

THIS WEEK

EDITORIALS

SPIDERS The amazing webs and weaving of our eight-legged friends **p.6**

WORLD VIEW The row over the Global Fund could spell disaster **p.7**



ILLUSION Perceived brightness is in the eye of the beholder **p.8**

The low-level nuclear threat

Europe is making a good start on learning about the health risks of low-dose radiation with a programme to share cold-war data and set research priorities. But the effort needs to be global.

Fear of the effects of an atomic strike haunted the politicians and scientists of the cold war. For years, researchers around the world worked on massive and systematic programmes to understand how ionizing radiation might affect survivors of a blast. Almost half a billion animals in the United States, the Soviet Union, Europe and Japan — mostly rats and mice, but also thousands of dogs and some rabbits and monkeys — were deliberately irradiated. These experiments were well designed and worked to identify the pathological consequences of doses of various types of radiation, delivered at different rates and by different routes, including inhalation or ingestion. Results were documented in detail and tissue samples were kept.

The political thaw of 1989 changed priorities, and these experiments and their results were largely forgotten. So were the scientists. But concern about ionizing radiation didn't fade entirely: instead, it was transferred increasingly to worries about low-level occupational exposure in the nuclear (and other) industries, and about exposure during medical diagnosis and treatment. The meltdown of the nuclear reactor in Chernobyl, now in Ukraine, in 1986 gave rise to new concern and political pressure, which was exacerbated by the Fukushima accident in Japan last year.

Science hasn't yet provided all the answers that governments need to respond to these concerns. Epidemiological studies, backed by animal experiments, have established beyond doubt that exposure to radiation levels above 100 millisieverts increases the risk of cancer in a predictable, dose-dependent way. But the risk to health at lower exposure levels is harder to pin down. In Germany, for example, the dose limit for occupational exposure is 20 millisieverts per year. But even when this limit is respected, there are more than 70,000 reports of suspected health damage among exposed workers each year, more than double the number in 1960. Some epidemiological studies suggest that low doses of radiation — as few as 10 millisieverts in children — may increase cancer risk in susceptible individuals, and may be associated with other conditions, such as heart disease and stroke. People living close to Fukushima are anticipated to receive around 10 millisieverts of accumulated radiation exposure each over the next decade.

Radiation-protection agencies want to check whether the weak associations between low-dose radiation and diseases are real. If they are, agencies need to understand the biological mechanisms behind them.

In 2008, the German Federal Office for Radiation Protection led an initiative with agencies from four other European countries and the European Commission to sketch out a rational, systematic plan to crack the low-dose problem. Their report was published in January 2009. It spawned the Multidisciplinary European Low Dose Initiative (MELODI), a platform of 15 European agencies and institutes that will develop and coordinate scientific policies. It also produced a 6-year, €13-million (US\$17.2-million) European Commission Network of Excellence to help to set up the platform and address some of its questions. Keen to maintain the musical theme, the network calls itself Low

Dose Research Towards Multidisciplinary Integration, or DoReMi.

Epidemiology has an important role in radiation research, but only when the dose received by an individual can be calculated with confidence, as in medical exposure. There must also be research on animals, posing questions not so different from those asked by radiobiologists during the cold war. Modern scientists won't be able to do animal

"Agencies want to check whether the associations between low-dose radiation and diseases are real."

experiments on the scale of those carried out before 1989, which would not be considered ethically justifiable by current standards. But they don't need to. Already, some scientists have started to analyse the cold-war data and materials, and have found that many of them should be usable.

In Rome last week, biologists from radiation-protection agencies and research institutes around the world who have taken up the challenge of low-dose radiation research met to discuss how to share contemporary information, and how to make the vast archives of historical data available to all.

It will involve a mammoth digitization and curation effort — expensive even before new experiments to plug gaps are factored in. Some members of MELODI estimate that it could cost up to €1 billion to understand the full biological effects of low-dose radiation. Yet the problem needs more than just money: it needs lobbying to attract biologists from other disciplines, who tend to dismiss radiobiology as old-fashioned.

MELODI has made a fine start on structuring a European scientific approach to a serious social problem. It now needs to orchestrate itself as a global platform — an unchained MELODI, if you like. ■

Poles apart

Protecting the Arctic from pollution requires a binding agreement like the Antarctic Treaty.

Last month's centenary of the ill-fated arrival of Robert Scott at the South Pole prompted a swell of events and publicity. Yet there is another noteworthy anniversary from the far south this month. Sixty years ago, Antarctica — perhaps for the first and only time — echoed to the sound of hostile gunfire. In what became known as the 1952 Hope Bay Incident, Argentinian soldiers fired a machine gun over the heads of a British landing party that was attempting to re-establish and supply an Antarctic base that had burnt down several years earlier. In response, the British despatched armed marines from the Falkland Islands, who forced the Argentinians to retreat

V. FLORESCU



Global health hits crisis point

The Global Fund's drive to ensure sustainability and efficiency means that it may not be able to meet its commitments to combat disease, says Laurie Garrett.

Last week, Michel Kazatchkine tendered his resignation as executive director of the Global Fund to Fight AIDS, Tuberculosis and Malaria. Regardless of whether you've heard of the French AIDS scientist, or even of the fund, you should keep reading. This is a crucial, dangerous moment for global health.

Kazatchkine made clear the political struggle that forced his resignation. "The Global Fund has helped to spearhead an entirely new framework of international development partnership," he wrote in his resignation letter. But under stress during the world economic crisis, with radically declining support from donors, a battle developed. "Today, the Global Fund stands at a cross-road. In the international political economy, power-balances are shifting and new alignments of countries and decision-making institutions are emerging or will have to be developed to achieve global goals. Within the area of global health, the emergency approaches of the past decade are giving way to concerns about how to ensure long-term sustainability, while at the same time, efficiency is becoming a dominant measure of success," he wrote.

It is almost possible to hear Kazatchkine spitting out the words 'sustainability' and 'efficiency'. Since the financial crisis of November 2008, a storm has been brewing over these concepts, one that affects everything from humanitarian responses to projects that distribute malaria bed nets. It is a fight, and on one side are those who believe that crises in general, and the AIDS pandemic and allied diseases in particular, constitute global 'emergencies' that must be tackled with full force, mistakes be damned. On the other are those who feel that AIDS is now a chronic disease that can be managed with medication and therefore requires investment in permanent infrastructure of care and treatment that can eventually be operated and funded by the countries themselves.

It is a classic battle of titans, pitting urgency against long-term sustainability. In his resignation letter, Kazatchkine essentially conceded victory to the forces for sustainability. Charitable urgency didn't stand a chance once the donor states started cinching their domestic budget belts so tightly that they had to punch new buckle holes.

The fund was established ten years ago as a unique mechanism to move billions of dollars from rich countries to poorer ones, to combat and treat three infectious diseases: HIV, malaria and tuberculosis. It acts as a granting agency, accepting applications from governments and health organizations, and convenes regular replenishment meetings to tell donors — mostly the governments of the United States, United Kingdom, France and Germany — how much money is needed for the next round.

By the end of 2009, the fund was disbursing US\$2.7 billion a year, and was underwriting almost half of all HIV treatment in poor

countries, about two-thirds of all malaria prevention and treatment in the world and about 65% of all tuberculosis efforts. The fund's most marked impact has been on malaria. At the end of 2011, the World Health Organization estimated that the number of malaria deaths had fallen by one-quarter between 2000 and 2010.

But Global-Fund cash has spawned dependency and expectation among its recipients. Should it disappear, or radically diminish, countries would be hard-pressed to finance malaria and tuberculosis efforts.

Indeed, the great diminishment has commenced. In October 2010, the fund asked donors for \$20 billion for five years' worth of disbursements. The donors were indignant and committed just over half that. In response, the fund's flabbergasted leadership cancelled the next grant round, and it will now not distribute new grants until 2014.

Donor scrutiny increased and a high-level independent review panel set up by the fund's governing board, which includes representatives of United Nations agencies and the World Bank, released a scathing report, citing a litany of problems, including fraud, theft and inconsistent decision-making by grant reviewers.

At a meeting in Accra, Ghana, on 21 November, the board members expressed shock at the problems identified by the high-level panel, and by reports commissioned on the situation on the ground in some countries. Some African leaders described riots and demonstrations at the lack of vital medicines, especially for HIV. The board's own investigation showed that the fund had committed assets of \$10 billion for 2011–13, but had

only about \$4 billion in its bank accounts.

The board called for ways to stretch available resources and eliminate inefficiencies. Key to that would be the appointment of a general manager to oversee all spending, pushing Kazatchkine aside. Stepping into that position is Colombian banker Gabriel Jaramillo.

To try to give Jaramillo a running start, in Davos, Switzerland, last week, Bill Gates handed over some \$750 million, redeemable by the fund in full during 2012, or spread out over time. And the Saudi Arabian government announced a \$25-million donation. As generous as these millions may be, the fund needs billions just to stay alive and fulfil country grants, let alone to grow. Right now we have no idea where that money will come from. Should the fund collapse, the consequences will be severe. Progress against tuberculosis and malaria will stall, and more than a million people living with HIV could be left without treatment. ■

Laurie Garrett is senior fellow for global health at the Council on Foreign Relations in New York, recipient of the 1996 Pulitzer Prize in Journalism and author most recently of *I Heard The Sirens Scream: How Americans Responded to the 9/11 and Anthrax Attacks*. e-mail: lgarrett@cfr.org

GLOBAL-FUND CASH
HAS SPAWNED
DEPENDENCY
AND EXPECTATION
AMONG ITS
RECIPIENTS.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/ge9til

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

CANCER GENETICS

Genome shatters in brain cancer

Cancer is usually attributed to a slow accumulation of genomic changes, but a few cancers result instead from a single catastrophic event that causes massive reshuffling of the genome. Researchers have discovered these major changes, called chromothripsis, in a type of medulloblastoma — a common childhood brain cancer — and have linked the disease to mutations in the tumour-suppressor gene *TP53*, which encodes the protein p53.

Jan Korbel at the European Molecular Biology Laboratory in Heidelberg, Germany, and his colleagues sequenced the genome of one patient with medulloblastoma and mutated *TP53*, and found many large genomic rearrangements. By analysing the genomes of 98 other patients with the brain cancer, they uncovered a strong association between mutant *TP53* and chromothripsis in one subtype of the cancer.

The authors propose that the protein p53 is involved in either initiating this massive genomic storm or keeping the cells alive afterwards.

Cell 148, 59–71 (2012)

PSYCHOLOGY

'Brightness' fools the eye

The pupils of the human eye shrink in response to brightness, even when the glow is merely an optical illusion.



Bruno Laeng and Tor Endestad at the University of Oslo used infrared eye trackers to monitor the pupils of participants looking at illusions of lightness or brightness. These graphic designs give the impression of having brighter or whiter components (pictured right) than similar images of equal luminance (left). The pupils rapidly constricted when people glanced at the 'brighter' image, then slowly readjusted to the picture's true light intensity.

The authors suggest that the physical response to the illusions may be an adaptation designed to

protect the eye's sensitive light-absorbing cells from potentially damaging levels of light.

Proc. Natl Acad. Sci. USA
<http://dx.doi.org/10.1073/pnas.1118298109> (2012)

EVOLUTION

Why animals get bigger over time

In palaeontology, Cope's rule holds that species evolve larger body sizes over geological time. One possible explanation has been that competition favours bigger bodies. To test this, Pasquale Raia at the University of Naples Federico II in Italy and his colleagues compiled a species tree of 554 extinct mammals across the past 60 million years, and analysed

size evolution within lineages.

They found that body size tends to increase as animals develop more specialized diets confined to particular habitats. Moreover, the origination of larger sizes coincided with periods of global cooling, and came at the cost of increased extinction risk.

Am. Nat. <http://dx.doi.org/10.1086/664081> (2012)

GENE THERAPY

An eye for gene repair

Gene therapy in dogs can reverse retinal defects that lead to blindness in humans.

William Beltran and Gustavo Aguirre at the University of Pennsylvania in Philadelphia



RESTORATION ECOLOGY

New wetlands don't measure up

Wetland restoration may be falling short of its goals, with restored or created wetlands lagging behind reference ones in terms of carbon storage and native species richness and abundance.

In many parts of the world, humans have destroyed more than half of the wetlands and efforts to restore them (pictured) stretch back 60 years. David Moreno-Mateos at the University of California, Berkeley, and his colleagues

analysed 621 restored or created wetlands. They found that carbon storage in such wetlands was just half that of reference wetlands even two decades after restoration. Moreover, restored and created wetlands showed only a 74% recovery in a measure of 'biological structure' that combined several measures of the number and richness of native species.

PLoS Biol. 10, e1001247 (2012)

BILL HUBBICK

US NATL ACAD. SCI.

and their group targeted a form of retinitis pigmentosa. This is a common cause of blindness in which mutations in the gene *RPGR* lead to the death of light-sensing photoreceptors in the retina. The authors tested their gene-therapy approach in four dogs with mutations that model a form of the disorder. After injecting a virus carrying human *RPGR* into one eye of each animal, the authors found that the photoreceptors' structure appeared normal, and that their light-responsiveness had improved in three of the four treated eyes. In the untreated eyes, the cells were shrunken, fewer in number and less responsive to light.

The authors say that their findings open the door to gene therapy in humans for certain kinds of retinitis pigmentosa. *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1118847109> (2012)

PHYSICS

Quantum motion captured

Even if cooled to a temperature of absolute zero, all objects will retain a fundamental jitter in their physical positions due to quantum 'zero-point' fluctuations. Oskar Painter and his colleagues at the California Institute of Technology in Pasadena have, for the first time, detected this motion in a tangible object and measured an asymmetric property that arises from it.

The researchers patterned a one-micrometre-wide silicon beam to allow for a strong interaction between laser light sent into the beam and the motion of the beam. As the beam was cooled to below 1 Kelvin, a striking asymmetry emerged. The energy of the laser photons increased the motion of the beam, but, as per the laws of quantum mechanics, the beam's fundamental jitter could not add to the energy of the laser photons.

Phys. Rev. Lett. 108, 033602 (2012)

GENOMIC MEDICINE

Sequencing diagnoses disease

In a study that shows both its promise and limitations in the clinic, researchers have used genomic sequencing to diagnose the genetic cause of disease in individual children — but succeeded in only a small proportion of them.

A team led by David Thorburn at the Murdoch Childrens Research Institute in Melbourne, Australia, and Vamsi Mootha at Harvard Medical School in Boston, Massachusetts, sequenced DNA from 42 unrelated children with symptoms of inherited disorders of the mitochondria — the cell's energy-producing organelles. The authors sequenced the mitochondrial DNA, as well as the coding regions of more than 1,000 nuclear genes that encode mitochondrial proteins. They found mutations in known or possible disease-causing genes in 23 of the children, and 10 received a firm diagnosis.

In the other cases, the causative gene may fall outside the sequenced regions, or the disease may be caused by weak interactions between several mutated genes, the authors say. *Sci. Transl. Med.* 4, 118ra10 (2012)

CANCER

Drug drives cancer stem cells

Cancer drugs that attack tumour-sustaining blood vessels may spur proliferation of the stem cells that contribute to the disease.

One such drug, bevacizumab, fails to prolong the survival of patients with advanced breast cancers, and studies have shown that patients on similar drugs often relapse. To find out why, Max Wicha and his team at the University of Michigan in Ann Arbor implanted human breast tumours into mice and treated them with bevacizumab and another

COMMUNITY CHOICE

The most viewed papers in science

ANIMAL BEHAVIOUR

Hair defence against bed bugs

HIGHLY READ
on <http://royalsocietypublishing.org> in December

Hirsute individuals are better able to sense blood-sucking bed bugs crawling on their skin than are their less hairy counterparts.

Isabelle Dean and Michael Siva-Jothy at the University of Sheffield, UK, tested the ability of 29 volunteers to detect the presence of bed bugs (*Cimex lectularius*) on a shaved and an unshaved arm. The volunteers detected the insects more frequently with their unshaved arm than their shaved one, and unshaved arms with a greater density of hair were more sensitive to the insects. Hairier skin also upped the time bed bugs took to find a feeding location.

Humans may have evolved less body hair to make parasites easier to see and remove, but the authors suggest that fine body hair has been maintained to enhance their detection. *Biol. Lett.* <http://dx.doi.org/10.1098/rsbl.2011.0987> (2011)

blood-vessel-blocking drug, sunitinib. Treated tumours produced more cancer stem cells than untreated ones. The low-oxygen environment that the drugs create in tumours switches on a molecular pathway that encourages cancer stem cells to divide.

Combining these drugs with others that target cancer stem cells could yield a better outcome, the authors suggest.

Proc. Natl Acad. Sci. USA <http://dx.doi.org/10.1073/pnas.1018866109> (2012)

GEOLOGY

Explosion in Death Valley

The volcanic eruption responsible for the giant Ubehebe Crater in California's Death Valley may have occurred more recently than previously thought. This could mean that the risk of similar explosions happening today is higher than anticipated.

When magma encounters groundwater, distinctive explosions, such as those of Ubehebe, ensue. But how this could have happened in Death Valley, the driest area in North America, has been unclear.

Brent Goehring at Purdue University in West Lafayette, Indiana, and his colleagues dated the crater (pictured) using the rate of accumulation of beryllium-10 isotopes in rocks exposed by the eruption. They conclude that the crater probably appeared 800–2,100 years ago, during a period in which Death Valley underwent prolonged drought.

The authors suggest that there may be enough permanent groundwater to drive further explosive eruptions in Death Valley. *Geophys. Res. Lett.* <http://dx.doi.org/10.1029/2011gl050130> (2012)



OCEAN/CORBIS

➔ **NATURE.COM**

For the latest research published by Nature visit: www.nature.com/latestresearch

SEVEN DAYS

The news in brief

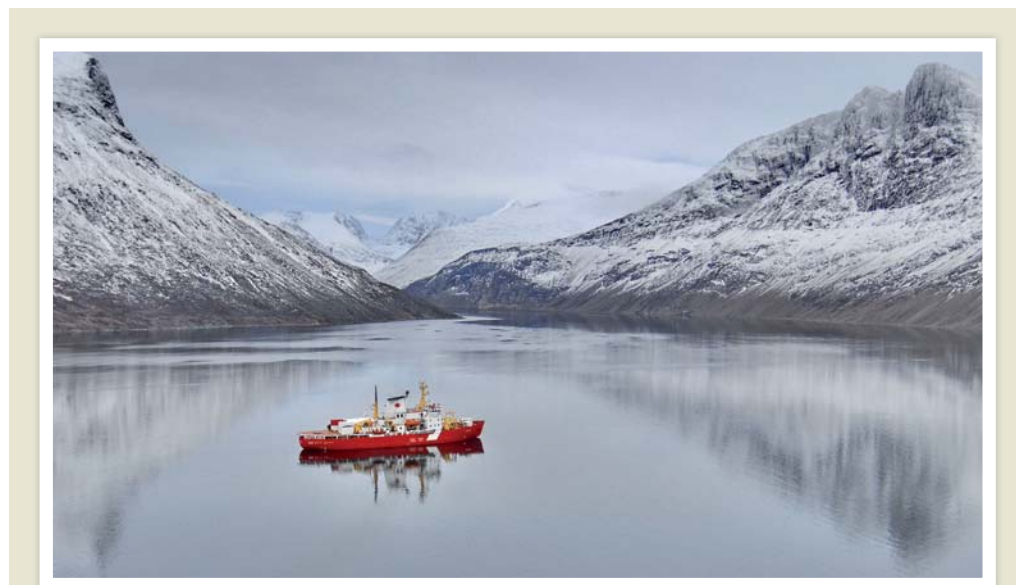
POLICY

Global Fund turmoil

The executive director of the Global Fund to Fight AIDS, Tuberculosis and Malaria, Michel Kazatchkine, has resigned after five years in charge. Kazatchkine, a French clinical immunologist, said on 24 January that he couldn't continue after the fund's board created the position of general manager, part of a restructuring plan agreed after the fund found corruption affecting some US\$39 million of its grants (see *Nature* **470**, 6; 2011). Based in Geneva, Switzerland, the fund is struggling to raise cash and has frozen the funding of new grants until 2014. But on 26 January, it received a \$750-million bonus pledge from the Bill & Melinda Gates Foundation. See page 7 for more.

Russian restraints

Physicists at one of Russia's leading research institutes have mounted an Internet campaign to draw attention to strict new budgetary and security policies that they say are hindering their work. Last year, Russia's government moved the Institute for Theoretical and Experimental Physics (ITEP) in Moscow, and two other



K. LEVESQUE/ARCTICNET

Engine trouble scuppers Arctic work

Canada's Arctic research vessel the CCGS *Amundsen* (pictured) has been put out of action while cracks in four of its six engines are repaired, forcing researchers to cancel the ship's research programme for 2012. The *Amundsen* can be used for science for up to 152 days a year, but only two six-week legs

had been planned for this year. "If it had to happen, this was a good year for it," says Martin Fortier, executive director of ArcticNet, a research network based at Laval University in Quebec City, Canada, that provides funding and logistics for Arctic research in Canada. See go.nature.com/7dqkt1 for more.

institutions, out of the state-owned nuclear corporation Rosatom and into the jurisdiction of the Kurchatov Institute in Moscow. The new administration has restricted trips abroad for scientists, and blocked visits from foreign researchers, ITEP physicists say. See go.nature.com/getmp7 for more.

Nanotech risks

The US National Research Council has added its voice to calls for high-quality research into the risks of nanotechnology. A report released on 25 January urges a more coordinated research strategy to cover questions such as how many nanoparticles of different kinds are being released into the environment, and who is

being exposed to them. It also criticizes the dual role of the US National Nanotechnology Initiative: promotional activities should be separate from the oversight of research into risks, the report says. See go.nature.com/8zospl for more.

Indian space fracas

The former chairman of the Indian Space Research Organisation (ISRO) in Bangalore and three of his former ISRO colleagues have been barred from holding any government job. Gopalan Madhavan Nair was sanctioned after an investigation into a satellite deal between Antrix, the ISRO's commercial arm, and Bangalore-based telecommunications company, Devas Multimedia. The 2005

deal was well below market value, and the government scrapped it in February 2011. Nair and others deny wrongdoing; Nair blames his troubles on his successor at the ISRO, Koppilli Radhakrishnan. See go.nature.com/wl5avm for more.

Yucca's successor

The United States should create a programme to identify permanent and interim storage systems for nuclear waste, according to a 26 January report by a White House commission. In its final report, the commission — established in January 2010 after the administration of President Barack Obama halted work on the controversial proposed repository at Yucca Mountain

JOURNALISM AWARD



Nature reporter Ewen Callaway last week won a UK Medical Journalists' Association award for 'Fighting for a cause', a News Feature on chronic fatigue syndrome and viruses (*Nature* **471**, 282–285; 2011).

NASA/JPL-CALTECH/CORNELL/ARIZONA STATE UNIV.
in Nevada — recommended setting up a new agency with dedicated funding. It also proposed a “consent-based approach” to nuclear storage, which involves finding communities that are willing to host waste.

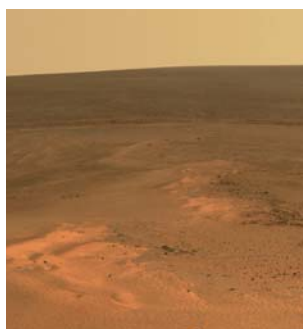
Japanese merger

Japan's government is planning to cut costs by merging five prominent research organizations. Plans announced on 20 January would create one body to oversee the RIKEN network of basic-research laboratories; the National Institute for Materials Science; the Japan Agency for Marine-Earth Science and Technology; the National Research Institute for Earth Science and Disaster Prevention; and the Japan Science and Technology Agency, the national funding body. See page 19 for more.

RESEARCH

Eight years on Mars

NASA's Mars rover Opportunity marked its eighth anniversary on the red planet on 25 January, perched on an outcrop called Greeley Haven (a view from which is **pictured**), on the rim of a massive crater named Endeavour. The rover, heading into its fifth Martian winter, has driven 34.4 kilometres so far. Its instruments are



fading: the decay of a cobalt-57 radiation source in the rover's Mössbauer spectrometer, for example, means that measurements of iron-containing minerals that used to take less than an hour now require weeks of work.

Planetary plethora

NASA's Kepler mission announced the discovery of 26 planets on 26 January, bringing the number of planets verified by the space probe to 61. Around 2,300 candidate planets still await confirmation. The latest findings are scattered across 11 planetary systems. See go.nature.com/hn3icx for more.

Top climate risks

Severe flooding is the United Kingdom's most pressing climate-change risk, according to an assessment of the nation's threats and opportunities under various climate scenarios. The study,

which was published on 25 January, has been lauded for its innovative methodology, although it acknowledges that its confidence in many of the findings is ‘low’ or ‘medium’. Britain will decide what actions to take when it publishes a national adaptation plan in 2013. See go.nature.com/7smiq5 for more.

PEOPLE

Chief scientist gap

The post of chief scientist at the US National Oceanic and Atmospheric Administration in Washington DC is to remain vacant after the administration of President Barack Obama withdrew its nomination of marine geochemist Scott Doney on 24 January — almost 18 months after Doney was put forward for the role. US Senator David Vitter (Republican, Louisiana) had put a hold on the confirmation process in protest at the administration's decision to issue a drilling moratorium after the April 2010 Deepwater Horizon oil spill in the Gulf of Mexico. The agency has not had a chief scientist since 1996.

FUNDING

Neglected diseases

Drug companies, philanthropists, national governments and the World

COMING UP

9 FEBRUARY

The maiden voyage of Europe's Vega rocket, a launcher for small satellites including research missions, begins. See go.nature.com/7smiq5 for more.

Health Organization (WHO) are coordinating to tackle neglected tropical diseases (NTDs) such as sleeping sickness, Chagas disease, leprosy and river blindness. In a joint announcement on 30 January, the WHO laid out a road map to control or eliminate ten NTDs by 2020, and drug companies announced expansions and extensions to drug-donation programmes amounting to 1.4 billion treatments a year. The Bill & Melinda Gates Foundation says that it will give US\$363 million over five years for NTD research. See go.nature.com/ksscgb for more.

BUSINESS

Illumina bid

A US\$5.7-billion takeover bid for the dominant company in the DNA-sequencing industry spurred a rally in genome-technology stocks last week. On 25 January, the drug giant Roche, based in Basel, Switzerland, announced its intention to buy Illumina in San Diego, California. Roche is betting that the sequencing market will soon lead to clinical applications in diagnostics and drug discovery. Illumina's board of directors is still considering the proposal but has said it may offer investors cheap shares in the event that Roche starts to gain too much stock, in an effort to make a takeover more expensive. See go.nature.com/2vmedh for more.

NATURE.COM

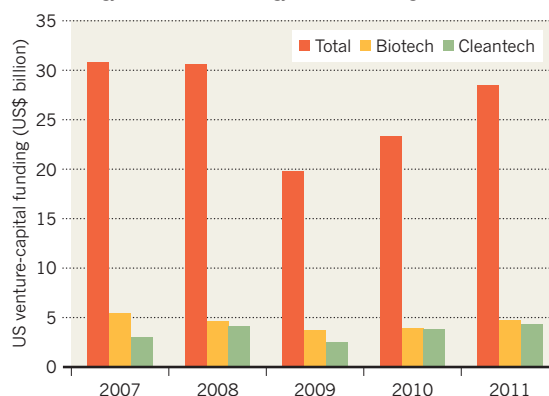
For daily news updates see: www.nature.com/news

TREND WATCH

Venture-capital biotechnology funding in the United States grew by 22% from 2010 to reach US\$4.7 billion last year, according to a 20 January report. Venture-capital funding rose by 22% overall (see chart), with clean technology up 12%. But investors are sticking to established firms: the number of biotech and medical-device firms raising their first round of venture-capital reached a 15-year low, says Jonathan Leff, managing director of private-equity firm Warburg Pincus in New York.

VENTURE CAPITAL'S RECOVERY

Investment in young US firms is nearing pre-recession levels, with the biotechnology and clean-technology sectors tracking the overall rise.



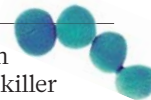
NEWS IN FOCUS

EMBRYONIC STEM CELLS Future of Californian funding agency is in doubt **p.15**

PHYSICS Subtle properties of the neutron could hold clues to a new theory **p.17**

JAPAN Leading geneticist fires parting shots as he quits his native country **p.18**

VACCINES A man with a plan to vanquish killer *Staphylococcus* **p.23**



B&C ALEXANDER/ARCTICPHOTO



Spills from pipes or rigs in icy areas such as those in Alaska are particularly challenging to clean up.

ENVIRONMENT

The great Arctic oil race begins

Conservationists fear spills in icy waters as Norway awards oil-production licences.

BY QUIRIN SCHIERMEIER
IN TROMSØ, NORWAY

The race is on for positions in the new oil provinces." That starting-gun quote was fired last week by Tim Dodson, executive vice-president of the Norwegian oil and gas company Statoil. The 'new oil provinces' are in the Arctic, which brims with untapped resources amounting to 90 billion barrels of oil, up to 50 trillion cubic metres of natural gas and 44 billion barrels of natural gas liquids, according to a 2008 estimate by the US Geological Survey.

That's about 13% of the world's technically recoverable oil, and up to 30% of its gas — and most of it is offshore.

Oil companies see an opportunity to sate the world's demand for fossil fuels. Green groups and many scientists, however, are horrified by the prospect of drilling and production in remote, often ice-choked waters, where spills would be harder to control and clean up than in warmer regions. Memories of the devastating environmental impact of the *Exxon Valdez* accident in 1989 in Alaska's Prince William Sound are still all too fresh — like the oil that can still be found in the area's beaches (see

Nature <http://doi.org/d3gqd2>; 2010).

At last week's Arctic Frontiers conference in Tromsø, Norway, the oil industry insisted that it will be cautious and responsible in extracting oil and gas in the region, and it rolled out an initiative to develop ways of coping with any accidents. Dodson told the meeting that "technology will be there to clean it up".

Statoil already operates the world's most northerly liquefied natural-gas production facility near Hammerfest, which draws gas equivalent to about 48,000 barrels of oil a day from the Snøhvit field in the Arctic waters off Norway. By 2020, the company hopes to extract one million barrels of oil equivalent a day from new wells in the Arctic. It is planning exploratory drilling later this year, for example, in the Skrugard and Havis gas fields that were discovered in the Barents Sea last year.

The Norwegian government is happy with Statoil's bold plans. Norway is currently the world's second-largest gas exporter, with production continuing to rise, but it is looking to the Arctic to offset a one-third decline in production at its oil fields farther south since 2000. "If we don't invest, we might lose another third within the next decade," says Ola Borten Moe, Norway's minister of petroleum and energy.

On 17 January, Moe awarded 26 production licences for developed offshore oil areas in the Norwegian and Barents Sea to companies including Statoil, Total, ExxonMobil and ConocoPhillips. And the settlement in 2010 of a long-running row between Norway and Russia over their Arctic maritime boundary will allow more exploration in formerly disputed parts of the Barents Sea (see 'Frozen fuels'). "There's an ocean of new opportunities that we will grasp with both hands," says Moe.

The resource rush is alarming critics. A group of 573 scientists, for example, wrote last week to US President Barack Obama, urging caution in authorizing new oil and gas activity in the Arctic Ocean north of Alaska. The open letter, coordinated by the Pew Environment Group, a conservation organization headquartered in Washington DC, argues that more research is needed to assess the potential impact on the region's environment and

ecosystems before going ahead with more drilling.

The industry holds that Arctic oil and gas development can be done in an environmentally ►

➔ **NATURE.COM**
Read more at
Nature's Arctic
special:
go.nature.com/s5w2pr

► sustainable manner despite the challenges. “We realize that there are huge issues when working in the cold and darkness and in the presence of sea ice in areas at great distance from any infrastructure,” says Joseph Mullin, a London-based programme manager at the International Association of Oil and Gas Producers. Mullin will oversee a four-year, US\$20-million research programme to address those issues, launched at the Tromsø conference by nine major oil companies.

The initiative, which is open to academic collaborators, will include research on the environmental effects of Arctic oil spills, spill trajectory modelling and remote sensing, and oil recovery techniques in sea-ice areas. It will also test Arctic clean-up technologies in a number of controlled oil releases. “You’d like to have a variety of spill-response options in the tool box before you venture out there,” says Mullin.

The leading Russian oil and gas companies,



FROZEN FUELS

A 2010 Norwegian–Russian border agreement has opened up the Barents Sea to oil companies eager to exploit resources.

Oil and gas facilities: ● Planned ● Operational
■ Oil and gas reserves

Gazprom and Rosneft, have so far stayed clear of the initiative, adding to concerns about their compliance with national and international safety standards.

In December 2011, for example, at least 37 people were killed when an oil rig under contract to Gazprom capsized off Sakhalin Island in Russia’s Arctic Ocean, resulting in a fine for the company. And according to Vladimir Chuprov, a Moscow-based energy expert who works for Greenpeace, emergency contingency plans for the Prirazlomnoye oil platform in the Russian Barents Sea, where commercial drilling is to start this year, have not been publicly released, despite being required by Russian regulators.

But even companies with better safety records should avoid the Arctic, say Chuprov and other environmentalists. “In our view no company is ready for offshore oil projects in the Arctic Ocean,” he says. ■

PSYCHIATRY

Diagnostics tome comes under fire

Field tests of new criteria are flawed, critics argue.

BY HEIDI LEDFORD

Diagnoses of certain mental illnesses could rise significantly from next year, say some mental-health experts — but not because of any real changes in prevalence. Instead, the critics blame what they say is a flawed approach to testing the latest version of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*, the standard reference used by researchers and mental-health professionals in the United States and many other countries

to assess patients, inform treatment, design studies and guide health insurers.

Changes to the diagnostic criteria in the fifth edition of the manual, *DSM-5*, due to be published in May 2013 by the American Psychiatric Association (APA) in Arlington, Virginia, have raised concerns that some disorders will be overdiagnosed (see table). Critics say that the analysis of field tests of the new criteria won’t settle those concerns.

Trials of *DSM-5* conducted at 11 academic centres were completed last October. In a

Commentary published in the *American Journal of Psychiatry* (H. C. Kraemer *et al.* *Am. J. Psychiatry* **169**, 13–15; 2012), members of the task force explained that the aim was not to focus on the frequency of a given diagnosis under the proposed *DSM-5* criteria compared with that under the previous criteria. Because there is no accepted prevalence for most psychiatric disorders, they argued, it would be impossible to tell whether a rise in diagnoses reflects a true increase in the sensitivity of the revised criteria or simply a rise in the number of false positives.

That raised the hackles of some researchers, who say that without such comparisons it will be impossible to flag up the possibility that some categories will show an increased prevalence. “It’s a real step back,” says Thomas Widiger, a psychologist at the University of Kentucky in Lexington, who notes that trials of *DSM-IV* were careful to compare old and new diagnostic criteria to see which performed better.

Allen Frances, emeritus professor of psychiatry at Duke University in Durham, North Carolina, led the 1994 *DSM-IV* revision and is an outspoken critic of *DSM-5*. Frances

CONTENTIOUS PROPOSALS FOR *DSM-5*

Changes to diagnoses of some mental illnesses are causing disquiet about the consequences for patients.

Disorder	Change	Rationale	Reason for controversy
Autism spectrum disorder	Combines multiple <i>DSM-IV</i> diagnoses into one; changes number of criteria needed for diagnosis	Improved specificity of diagnosis	Concerns about underdiagnosis, loss of eligibility for state support
Attenuated psychosis syndrome	New diagnosis	To identify young people at risk for later manifestation of psychotic disorder	Concerns about overdiagnosis/overmedication
Major depressive episode	Removes “bereavement” exclusion	Evidence does not support separating loss of loved one from other stressors	Concerns about overdiagnosis/overmedication
Mild neurocognitive disorder	New diagnosis	To facilitate earlier intervention for individuals in early stages of neurocognitive disorders	Concerns about overdiagnosis/overmedication

FUNDING

Stem-cell agency faces budget dilemma

The California Institute for Regenerative Medicine plans for a future without state support.

BY ERIKA CHECK HAYDEN

acknowledges that the field trials for *DSM-IV* were far from perfect. For example, his trials failed to identify the dramatic surge in diagnoses of attention-deficit/hyperactivity disorder that followed changes made in *DSM-IV*. The trials suggested that there would be an increase of about 15% in the disorder. Instead, says Frances, the diagnosis rose threefold. “We missed the boat,” he says. “But at least we had some sense that there would be an increase.”

Results from the *DSM-5* academic field trials have yet to be presented, but early calculations suggest that, in general, there will be no big differences in the frequency of diagnoses, says Darrel Regier, vice-chair of the *DSM-5* task force and APA director of research. That claim has done little to alleviate concerns, however, because the trials enrolled patients who were initially diagnosed under *DSM-IV* standards. This leaves untested the possibility that the *DSM-5* criteria will capture many more patients who were previously deemed healthy, notes Widiger.

Observers are also alarmed by the statistical thresholds that the trials used to assess reliability, or the likelihood that two or more clinicians would arrive at the same diagnosis using the proposed criteria. This likelihood is often expressed as a statistical term called ‘Cohen’s kappa’. A kappa of 0 means that there is no agreement between the clinicians; a value of 1 means that the clinicians agree totally.

Researchers in the field often strive to reach a kappa of 0.6–0.8, indicating that the independent diagnoses agree more often than not. But in the Commentary, lead author Helena Kraemer, an emeritus statistician at Stanford School of Medicine in California, argued that a kappa of 0.2–0.4 could sometimes be acceptable. Kraemer later elaborated to *Nature* that the task force was largely aiming for a kappa of 0.4–0.6, but that it wanted to prepare the field for seeing values as low as 0.2 in particularly rare diagnoses or in those without biological markers.

Unlike tests on the previous edition, the reliability tests on *DSM-5* were performed on separate occasions, so that the clinicians involved were unaware of each other’s diagnoses. Widiger says that he supports the more rigorous approach, but that accepting a value as low as 0.2 gives him pause. “I’ve never seen anybody argue that a kappa of 0.2 is acceptable,” he says. “You just can’t get much lower than that.”

Not everyone is worried about a surge in diagnoses. Thomas Frazier, a paediatric psychologist at the Cleveland Clinic in Ohio, has carried out his own study of *DSM-5* criteria for autism spectrum disorder. His results, published online last year (T. W. Frazier *et al.* *J. Am. Acad. Child Adolesc. Psychiatry* 51, 28–40; 2012), suggested that the new definition would omit some patients with autism, but that this could be easily corrected by requiring one less symptom to meet the threshold for a positive diagnosis. “Unfortunately, the *DSM* committees are not systematically doing these kinds of studies,” he says. ■

Halfway through its initial ten-year mandate, the California Institute for Regenerative Medicine (CIRM) in San Francisco is confronting a topic familiar to anyone at middle age: its own mortality.

The publicly funded institute, one of the world’s largest supporters of stem-cell research, was born from a state referendum in 2004. Endorsements from celebrities such as then-state governor Arnold Schwarzenegger and the late actor Christopher Reeve, who had been paralysed by a spinal injury, helped to garner voter support for a public bond to underwrite the institute. But with half of the US\$3 billion that it received from the state now spent and the rest expected to run out by 2021, CIRM is now actively planning for a future that may not include any further state support.

“It would be premature to even consider another bond measure at this time,” wrote Jonathan Thomas, CIRM’s chairman, in a draft of a transition plan requested by the state legislature. Thomas outlined the plan on 24 January at a public hearing held in San Francisco by the US Institute of Medicine, which CIRM has asked to review its operations.

Given that California is facing severe budget shortfalls, several billion dollars more for stem-cell science may strike residents as a luxury that they can ill afford. It may also prove difficult for CIRM’s supporters to point to any treatments that have emerged from the state’s investment. So far, the agency has funded only one clinical trial using embryonic stem cells, and that was halted by its sponsor, Geron of Menlo Park, California, last November.

Yet the institute has spent just over \$1 billion on new buildings and labs, basic research, training and translational research, often for projects that scientists say are crucial and would be difficult to get funded any other way. So the prospect of a future without CIRM is provoking unease. “It would be a very different landscape if CIRM were not

around,” says Howard Chang, a dermatologist and genome scientist at Stanford University in California.

Chang has a CIRM grant to examine epigenetics in human embryonic stem cells, and is part of another CIRM-funded team that is preparing a developmental regulatory protein for use as a regenerative therapy. Both projects would be difficult to continue without the agency, he says. Federal funding for research using human embryonic stem cells remains controversial, and could dry up altogether after the next presidential election (see *Nature* 481, 421–423; 2012). And neither of Chang’s other

“It would be a very different landscape if CIRM were not around.”

funders — the US National Institutes of Health (NIH) and the Howard Hughes Medical Institute in Chevy Chase, Maryland — supports his interdisciplinary

translational work. Irina Conboy, a stem-cell engineer at the University of California, Berkeley, who draws half of her lab’s funding from CIRM, agrees that in supporting work that has specific clinical goals, the agency occupies a niche that will not easily be filled by basic-research funders. “The NIH might say that the work does not have a strong theoretical component, so you’re not learning anything new,” she says.

CIRM is developing plans to help its grantees to continue their work if the agency closes. One option is a non-profit ‘venture philanthropy’ fund that would raise money from private sources to support stem-cell research. The agency is also writing a strategic plan for the rest of its ten-year mandate that focuses on translating research into the clinic, acknowledging that CIRM’s best shot at survival — and at sustaining future funding for stem-cell researchers — could come from a clinical success.

As CIRM board member Claire Pomeroy, chief executive of the University of California, Davis, Health System in Sacramento, noted at the agency’s board meeting on 17 January: “If you asked the public what they would define as success, they would say a patient benefited.” ■

► **NATURE.COM**
Read more about
CIRM’s history and
impact:
go.nature.com/bpeds

MEDICINE

Informed consent on trial

Lengthy, complicated documents leave many clinical-trial participants in the dark about the risks they face.

BY DANIEL CRESSEY IN BRUSSELS

Informing clinical-trial participants of the risks they face is a cornerstone of modern medical research, and it is enshrined as a human right in international codes of ethics. But an influential group of ethicists and medical researchers warned at a meeting in Brussels last week that the process has become a box-ticking exercise focused more on offering legal protection to a trial's organizer than actually protecting patients.

"We clearly identified there is an urgent need to do informed consent better," says Ingrid Klingmann, chairman of the European Forum for Good Clinical Practice, a think tank based in Brussels and the organizer of the meeting. "The pressure is really huge on all those involved to better enable patients to understand the implications of their study participation, their benefits, risks and obligations."

The European Clinical Trials Directive, which sets minimum standards for clinical trials conducted in the European Union's member states, says that trial participants must be duly informed of the "nature, significance, implications and risks" of the clinical trial. Yet delegates at the meeting detailed a host of ways in which the system fails to meet those criteria.

Denis Lacombe, director of the Brussels headquarters of the European Organisation for Research and Treatment of Cancer, presented data showing that the median length of consent forms used in the organization's trials tripled between 1995 and 2009 (see 'Fatter forms'). In many cases "the informed consent is 20 pages of nothing," says Harry Bleiberg, an oncologist formerly at the Jules Bordet Institute in Brussels who is now a medical consultant for the pharmaceutical industry.

Other speakers showed that warnings of possible side effects are often vague, and jumble potentially serious problems among

minor concerns. Details of the prognosis of the disease and the likely outcome of the treatment are often scarce.

Long, jargon forms mean that many patients have little idea what they are actually signing up to, says Dominique Sprumont, a health lawyer at the University of Neuchâtel in Switzerland. In a study of about 200 patients involved in cancer clinical trials in the United States who responded to a questionnaire, 63% were unaware of the additional risk associated with participating, and 70% did not realize that the treatments being offered were unproven (S. Joffe *et al. Lancet* 358, 1772–1777; 2001).

FATTER FORMS

Europe's cancer-trial consent forms are much longer — but are they more informative?



Data from protocols of 248 trials run by the European Organisation for Research and Treatment of Cancer.

One solution presented at the meeting was to base forms on those used in Norway, where a short summary of the main aspects of a trial precedes more detailed information. And many delegates advocated a greater use of graphics and charts. "We will certainly work on a proposal for a template that includes a one- or two-page summary," says Klingmann, who hopes to circulate the document to clinical-trial organizers around the world.

Some ethicists advocate digital consent forms, delivered by computer or smart phone. Interactive forms could lead participants through the consent process, and keep them more informed during the trial. They could also make it easier for researchers to obtain new consent if they need to change a study once it is under way.

GLOBAL ISSUE

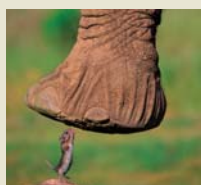
Concern over the issue is not unique to Europe. Trials are increasingly 'off-shored' to less-developed countries, and several investigations have found poor consent procedures there. Last month, for example, GlaxoSmith-Kline was fined in Argentina over what the company describes as "administrative irregularities in the process of obtaining informed consent from a small proportion of patients" in a trial in which 14 children had died (the company is appealing the ruling).

The US Department of Health and Human Services is currently consulting about revising clinical-trial regulations, and looks likely to demand shorter, more understandable consent forms, although there is no specific deadline for a decision. "Our effort is to enable a patient who is contemplating enrolling in a clinical trial to have clear and full information to use in making a decision," says Ann Bradley, a spokeswoman for the department's Office for Human Research Protections. A number of US medical groups are backing the effort: the American Society of Hematology in Washington DC, for example, says that in addition to making the process less burdensome for researchers, shorter forms could help to increase the number of patients recruited to clinical trials.

Sprumont argues that the scientists involved in clinical trials should also work to improve the system. "Scientists must assume more responsibility," he says. ■



TOP STORY



From mouse to elephant in a few (million) generations go.nature.com/fijfto

MORE NEWS

- Cancer drugs affect mouse genomes for generations go.nature.com/rdmmhy
- Spiders dodge cannibalism with remote copulation go.nature.com/6egpxl
- Visiting Iranian scientist falls foul of US authorities go.nature.com/lp71fw

PODCAST

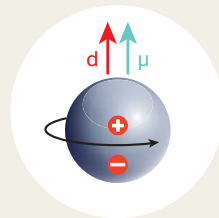
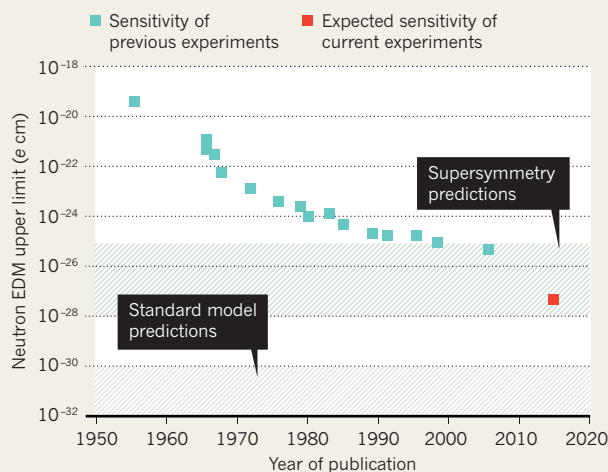


The dangers of sugar, predicting supervolcanoes and bringing elephants to Australia go.nature.com/xuyueh

SOURCE: A. KNECHT

LOWERING THE BAR

The standard model of particle physics predicts a small electric dipole moment (EDM) for the neutron, well below the sensitivity of previous experiments. A larger dipole, predicted by some versions of supersymmetry, should lie within the range of three current experiments.



If the internal structure of the neutron creates an electric dipole (d), its orientation with respect to magnetic spin (μ) would change under a reversal of charge and parity (CP), thus violating CP symmetry.

PHYSICS

Dipole hunt stuck in neutral

Physicists admit to delays as attempts to measure the neutron's charge are plagued by difficulties.

BY EUGENIE SAMUEL REICH

Discovered 80 years ago this month, the neutron is famous for what it lacks: electric charge. Yet this is a simplification of a deeper truth. Each neutron is made of charged quarks: two down quarks that each carry a negative charge one-third as strong as an electron's; and an up quark that carries a two-thirds positive charge.

In theory, this arrangement could create an electric dipole moment (EDM) — an offset between the centres of positive and negative charge. Measuring the EDM could reveal clues to physics as profound as that sought at the Large Hadron Collider (LHC) at CERN, Europe's particle-physics lab near Geneva, Switzerland. Unlike the LHC experiments, which rely on sheer energy, dipole measurements are so subtle and easily foiled that the three leading efforts are woefully behind schedule.

This week, a US team that is developing a neutron EDM experiment at the Oak Ridge National Laboratory in Tennessee is awaiting feedback from a technical-review committee after delays forced the group to update its research plans. A team at the Laue-Langevin Institute in Grenoble, France, had intended to

begin taking data in 2009, but is still running tests on its apparatus. A third group, based in Switzerland, is aiming to collect its first batch of data later this year after problems prevented it from beginning its run in 2011.

"The experiments have all taken a lot longer than any of us thought they would," says Philip Harris, a physicist at the University of Sussex, UK, and spokesman for the Grenoble-based effort.

The stakes are high. An EDM for the neutron would violate charge-parity symmetry, which dictates that particle interactions are unchanged when particles are replaced by their mirror images. The standard model of particle physics predicts subtle violations of charge-parity symmetry and proposes an EDM that is below the sensitivity of any currently achievable experiment. But a larger neutron EDM, possibly detectable by the latest efforts, could point to physics beyond the standard model (see 'Lowering the bar'). One popular theory, called supersymmetry — which predicts a zoo of undiscovered particles and interactions — includes the possibility that a process known as electroweak baryogenesis took place in the early Universe, producing more matter

than antimatter and breaking the symmetry between charge and parity.

Physicists at the LHC are also seeking hints of supersymmetry. Yet as the US\$6.5-billion LHC has forged ahead, the neutron EDM experiments, with modest budgets of tens of millions of dollars, have hit a wall. The Swiss-based experiment at the Paul Scherrer Institute in Villigen aims to measure the dipole moment by comparing the movement of neutrons in a strong electric field to that of mercury atoms in the same field. But the experiment has been sidetracked by contamination from an unknown source that destroyed the polarization of the mercury atoms. Researchers have had to replace several parts of the apparatus to solve the problem, says project spokesman Klaus Kirch, a physicist at the Swiss Federal Institute of Technology in Zurich.

The Grenoble-based experiment has run into different problems. It uses liquid helium to cool neutrons to almost absolute zero, then observes whether their behaviour changes in an electric field. But the researchers have been stymied because the liquid helium boils and begins to break down as stronger electric fields are applied.

The US experiment, co-funded by the Department of Energy and the National Science Foundation, has had a similar problem. It aims to detect the neutron's presumed EDM by looking for changes in the light emitted when cold neutrons interact with helium-3 nuclei in the presence of an electric field, a process that is subtly altered if the neutron has an EDM. The project's co-spokesman, Paul Huffman of North Carolina State University in Raleigh, says that trial runs on a test apparatus ran into problems when bubbles began to form on its electrodes. He says that the group is now searching for alternative electrode coatings and is considering increasing the pressure in the experiment, which might prevent the bubbling.

The US agencies funding that project are tracking the problem.

"The experiments have all taken a lot longer than any of us thought they would."

Last year, the Department of Energy delayed a decision on whether to approve the experiment by two years to allow for more exploratory research and develop-

ment. "What we basically did was to move the goalposts," says Peter Jacobs, a senior scientist at the Lawrence Berkeley National Lab in California, who was involved in the review. Now, a second technical-review committee is assessing the latest research plans for the experiment.

Despite the problems, nuclear and particle physicists continue to express broad support for the neutron EDM studies, which they say are a unique complement to the LHC work. "It's a constellation of experiments that is critical," says Michael Ramsey-Musolf, a theoretical physicist at the University of Wisconsin-Madison. ■

Genomics ace quits Japan

Yusuke Nakamura blames government inertia for his move to the United States.

BY DAVID CYRANOSKI IN TOKYO

As 2011 dawned, leading Japanese geneticist Yusuke Nakamura was launching a cabinet-level innovation office that he hoped would spur the country's stagnant biomedical industry. By the year's end he was preparing to leave for the University of Chicago in Illinois, frustrated by Japan's lack of support for genomics and doubtful that the country would ever take a global lead in medicine.

Nakamura has been a major force in Japanese science, with more than 1,150 publications to his name. He is one of the country's best-funded researchers and has led some of Japan's biggest science projects. In addition, he is a successful biotechnology entrepreneur, having established the company OncoTherapy Science, based in Kawasaki City.

But he has also been the target of criticism, both from opponents of 'big science' and from those who thought that his cancer-genetics research at the University of Tokyo put clinical science ahead of participants' rights¹ — a charge he rejects.

Some argue that Japanese science can shrug off Nakamura's departure. "He's not the only one that can do genomics," says Yuko Ito, a molecular biologist and science-policy researcher at the National Institute of Science and Technology Policy in Tokyo. "Genomics has gone far in many directions, been incorporated in so many other fields. It doesn't depend on one person."

But Tatsuhiko Shibata, head of genomics at the National Cancer Center Research Institute in Tokyo, says he is "very worried" by the move, adding that Japan has lost a strong lobbyist for the field. Without Nakamura's leadership, Shibata fears that Japan will fall further behind countries that are pressing ahead in genomics, such as China².

Some biologists still feel ashamed that Japan contributed only 6% of the human genome sequence published in 2001 (by contrast, the United States provided 59%). The country rebounded, largely as a result of Nakamura's efforts, by contributing about one-quarter of the International HapMap Project in 2005, which compared variations between human genomes³. But since then, money for genomics research in Japan has dried up. The country is leading just 1 of 50 projects under the International Cancer Genome Consortium, for example, compared with the United States' 20 and Mexico's 3.

Shibata says that genomics has lost out

because of Japan's shift to post-genomic research, which saw large-scale projects launched to investigate subjects such as protein structure⁴. Nakamura blames a more general rejection of "big science" and clinical applications by the "basic scientists" involved in science policy-making. "It's obvious that support

for genomics is not there," he says.



"By moving to the United States, I might be able to bring something back."

Speaking to *Nature* from his office at the University of Tokyo, Nakamura says that the loss of government support for genomics accounts for 20% of his decision to leave. The opportunities at Chicago, which recently opened a Center for Personalized Therapeutics that will incorporate genetic information into routine clinical

practice, contributed another 30%. And 10% came from a tussle with the newspaper *Asahi Shimbun* over the way a story linked him with a cancer-vaccine clinical trial in which a participant experienced an adverse reaction, he says. He is suing the paper and two journalists.

But the remaining 40% — the biggest single factor — is that he felt powerless at the government's new Office of Medical Innovation. A year ago, he hoped to resuscitate Japan's

moribund pharmaceutical and medical-device industries, and redirect its ballooning health-care outlays from foreign pharmaceutical coffers to domestic companies. Nakamura laments that none of the cancer drugs approved by the US Food and Drug Administration in the past decade came from Japan, and he has pushed for innovations such as a national high-throughput drug-screening facility⁵. He has also tried to persuade the government to integrate his plans into the economic stimulus programme that followed last year's earthquake and tsunami. But, he says, flatly, "my proposals were not taken".

He acknowledges that the government is investing in biomedicine as part of its quake-recovery plan. A ten-year, ¥64-billion (US\$835 million) initiative launched in November — tentatively named the Tohoku Medical Megabank — aims to create the world's largest collection of biological samples at Tohoku University, which is in the quake region. But Nakamura says he was left out of the planning. "The innovation office plays no role. I had nothing to do." To add to Nakamura's frustration, it is unclear how Megabank will link to the existing BioBank Japan, a project he initiated that now holds about 200,000 samples (see 'Japan plans another huge tissue repository').

University of Tokyo vice-president Yoichiro Matsumoto, a mechanical engineer who will take over Nakamura's role at the innovation office, shares his sense of crisis. "If we don't create an innovation system now, we'll be sunk," he says. He hopes to make the office useful

BIOBANK, MEET MEGABANK

Japan plans another huge tissue repository

Japan is moving to establish a massive patient-sample bank — the Tohoku Medical Megabank — using stimulus money provided to boost the economy and replace infrastructure in the northeastern region devastated by a tsunami last March.

Masayuki Yamamoto, a cancer biologist at Tohoku University who is expected to lead the programme, says that in the first three years samples will be gathered from 70,000 healthy people from about 10,000 three-generation families; and from another 80,000 unrelated people. This will provide a resource for studying the relationship between genes and lifestyle factors in

disease. The third cohort, to follow several years later, will target patients with specific diseases. How the Megabank will link in with the country's current repository, BioBank, remains to be seen.

The Megabank will be supported by medical facilities in the Tohoku region, where reconstruction plans include an electronic medical records system, which will help with the sharing of genetic and other data. Launched with ¥15.8 billion (US\$207 billion) from a supplementary budget passed in November, the Megabank is expected to receive roughly ¥5 billion per year for the next decade. **D.C.**

by creating networks of researchers to support national projects such as the Megabank. Although he credits Nakamura as a brilliant scientist, he adds that his strengths lie more in research than administration and networking.

Atsushi Sunami, a science-policy researcher at the National Graduate Institute for Policy Studies in Tokyo, blames Nakamura's departure on bureaucracy and a lack of political

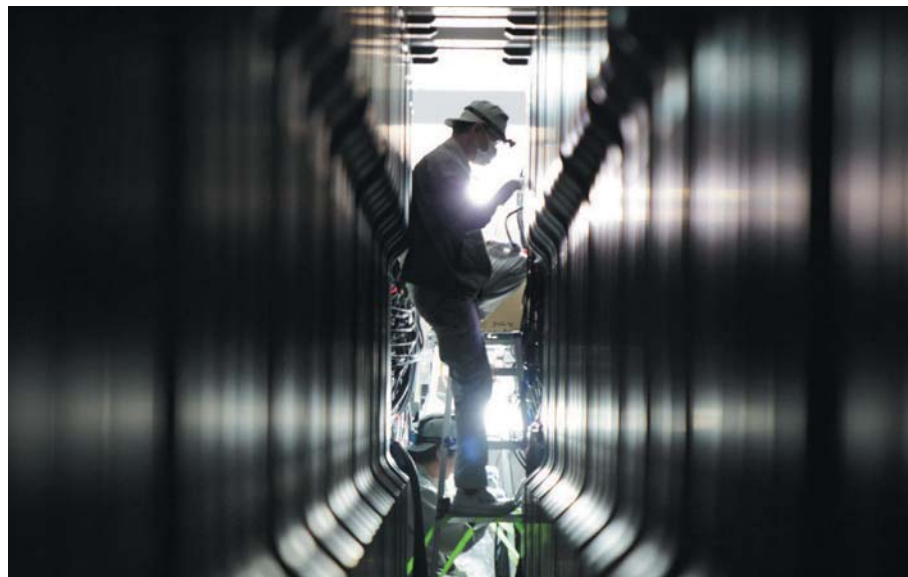
leadership, rather than a fundamental problem with medical innovation. "I think he just couldn't handle the system" of Japanese government, Sunami says.

When Nakamura returned to Japan in 1989 from a position at the Howard Hughes Medical Institute at the University of Utah in Salt Lake City, he says, his ambition "was to make drugs in Japan and provide them to Japanese patients".

But now, he says, "if I stay, I wouldn't be able to make anything. By moving to the United States, I might be able to bring something back." ■

1. Cyranoski, D. *Nature* **424**, 359 (2003).
2. Cyranoski, D. *Nature* **464**, 22–24 (2010).
3. The International HapMap Consortium *Nature* **437**, 1299–1320 (2005).
4. Cyranoski, D. *Nature* **443**, 382 (2006).
5. Cyranoski, D. *Nature* **474**, 136 (2011).

JUI PRESS/AFP/GETTY IMAGES



RIKEN, home to the world's fastest supercomputer (above), may be overseen by a new body.

RESEARCH POLICY

Japan plans to merge major science bodies

Drive to save money could increase bureaucracy.

BY ICHIKO FUYUNO

In its battle against a sluggish economy, Japan's government is gearing up to make cost savings through a root-and-branch reform of the country's science system, merging some of its most prominent research organizations.

Plans approved by the government's cabinet on 20 January would consolidate the RIKEN network of basic-research laboratories with the National Institute for Materials Science (NIMS), the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), the National Research Institute for Earth Science and Disaster Prevention (NIED) and the Japan Science and Technology Agency (JST) — the national funding body.

"This is the result of our aim to create the

world's leading science and technology while slashing wasteful spending and making the organization as compact as possible," said Tenzo Okumura, a senior vice-minister of the education and science ministry, at a press conference last week. The policy would probably create an overarching body to supervise all five institutions, which would share more of their research and administrative resources, and lose some of their executive directors.

But with few details about the timing, potential cost savings or full implications of the change, many researchers are concerned that it could be a recipe for harsh funding cuts and even greater bureaucracy.

"The plan isn't intended to improve the research system," says Tasuku

Honjo, a molecular biologist at Kyoto University and a former member of the Council for Science and Technology Policy, the government's top science advisory body. "Each of the institutes has fundamentally different characteristics, so I feel they would be very difficult to integrate effectively."

JST president Michiharu Nakamura, however, is optimistic about the change. "The new organization will be expected to play the main part in leading the national innovation," he says. "We will work hard to keep the independence of our fund-distribution process after the integration, as well as to strengthen the cooperation between stakeholders such as the academic community and industries." The plan would also merge the National Institute of Radiological Sciences (NIRS) with the Japan Atomic Energy Agency. The government says that this will speed up and strengthen its commitment to mitigating the Fukushima accident, and help to improve nuclear safety measures.

The news comes as the government scrambles to gain public support for tax hikes while cutting public spending. Since the Democratic Party of Japan took power in 2009, the Government Revitalization Unit (GRU), a budget watchdog chaired by the prime minister, has been reviewing government-funded projects and organizations to find savings (see *Nature* <http://doi.org/bs75dv>; 2010).

"The GRU rapidly came up with the plan a month ago and pushed forward without much discussion," says Atsushi Sunami, a science-policy specialist at the National Graduate Institute for Policy Studies (GRIPS) in Tokyo. Legislation to enact the overhaul will be submitted to Japan's parliament, the Diet, in the coming months. If passed, the changes would take effect in April 2014. More institutes may be added to the integration plan during discussions in the Diet, say policy experts.

The government also intends to set up a third-party audit system to evaluate administration and performance more strictly. The idea of more governance worries Sukekatsu Ushioda, president of NIMS. "Currently, the NIMS president has sufficient authority, so the decision-making process is fast. We at NIMS can often take new actions shortly after a board meeting," he says. "But if a higher management is added above NIMS, the decision-making would be slower and it would be difficult to take dynamic approaches." ■

➤ **NATURE.COM**
For more from
this region, see
Nature Japan:
natureasia.com/japan



ATTACK OF THE BLOBS

Jellyfish will bloom as ocean health declines, warn biologists. Are they already taking over?

BY MARK SCHROPE

Last summer, intrepid surfers flocked to Florida's east coast to ride the pounding swells spawned by a string of offshore hurricanes. But they were not prepared for a different kind of hazard washing towards shore — an invasion of stinging moon jellyfish, some of which reached the size of bicycle wheels. The swarms of gelatinous monsters grew so thick that they forced a Florida nuclear power plant to shut down temporarily out of concern that the jellies would clog its water-intake pipes.

Earlier in the year, similar invasions had

forced shut downs at power plants in Israel, Scotland and Japan. The gargantuan Nomura's jellyfish (*Nemopilema nomurai*) found in Japanese waters can weigh up to 200 kilograms and has plagued the region repeatedly in recent years, hampering fishing crews and even causing one boat to capsize. Jellyfish have destroyed stocks at fish farms in Tunisia and Ireland. And in the Mediterranean Sea and elsewhere, officials have built nets to keep out jelly swarms.

The jellyfish blooms seem to bear out warnings from some scientists and conservationists, who argue that humans are knocking marine ecosystems off balance, causing a massive

Blooms of giant Nomura's jellyfish (*Nemopilema nomurai*) have troubled Japanese fishing crews.

increase in the global population of jellyfish — a catch-all term that covers some 2,000 species of true cnidarian jellyfish, ctenophores (or comb jellies) and other floating creatures called tunicates. But many marine biologists are now questioning the idea that jellyfish have started to overrun the oceans.

This week, a group of researchers published preliminary results from what will be the most comprehensive review of jellyfish population data¹. They say that there is not yet enough evidence to support any conclusions about a global upswing in jellyfish populations. "We are not at a point to make these claims," says Robert Condon, a marine scientist at the Dauphin Island Sea Lab in Alabama, who leads the group. "If you look at how that paradigm formed, it's not based on data or rigorous analyses."

The main problem, say the scientists, is that jellyfish are notoriously difficult to study, so they have historically received scant attention from marine biologists. There are remarkably few data about their life cycles, populations and responses to natural oceanographic cycles. But the creatures could serve as key indicators for the health of the oceans, so scientists are now building a database of jellyfish research and exploring new ways to keep track of them.

Monty Graham, chairman of the department of marine science at the University of Southern Mississippi in Diamondhead, is part of the review team now questioning the idea of a rise in jellyfish numbers, but more than a decade ago he was sounding the alarm. In

KYODO/NEWS.COM

1996, he took a position at the Dauphin Island Sea Lab, where he found that the US National Oceanic and Atmospheric Administration had years of mostly unprocessed population data on moon jellyfish (*Aurelia aurita*) and Atlantic sea nettles (*Chrysaora quinquecirrha*) in the Gulf of Mexico. These data were a rare treasure; only a handful of similar long-term records exist.

Graham discovered that from 1985 to 1997, the jellies had grown substantially more widespread and abundant in several parts of the Gulf, and he suggested that human changes to the ecosystem might be the cause².

Similar findings supported the notion. In the Bering Sea, one of the handful of locations with a monitoring record longer than a few years, jelly numbers had also risen through the 1990s. That matched predictions made by ocean scientists, who had warned that as humans degrade the oceans they are shifting ecosystems, reducing numbers of larger fish and promoting populations of organisms from lower down the food chain^{3,4}. Among the beneficiaries would be algae, toxic plankton and jellyfish — in other words, there would be a sea of slime.

BOOM AND BUST

The link between ocean degradation and jellyfish makes biological sense. Nutrient pollution can increase food supplies for jellyfish; overfishing can reduce their competition; and warmer temperatures are thought to trigger reproduction in some jellyfish species.

But as the slime paradigm gained traction in the literature, something odd was happening in the Gulf of Mexico. One of Graham's graduate students, Kelly Robinson, found that jellyfish numbers in the northern Gulf had declined for several years after 1997, and then rebounded. Her work has not yet been published.

Researchers studying jellyfish in the Gulf of Mexico and the Bering Sea now think that long-term natural climate cycles have an important role in controlling populations there. "Just seeing a lot of jellyfish does not say anything," says Graham now. "People say, 'Oh my God, the world is going to hell,' but jellies form blooms. That's what they're supposed to do." The challenge for researchers lies in separating normal fluctuations from those for which humans might deserve some of the blame.

Graham and others decided to take a scientific step back. In 2009, the National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara, funded Graham, Condon and Carlos Duarte, a marine ecologist at the Mediterranean Institute for Advanced Studies on Majorca, Spain, to establish the working group that has just published its findings. Made up of dozens of researchers, it compiled all the scientific data available on jellies worldwide. After a preliminary examination, the group said¹ that it could not support the conclusion that jellyfish numbers are increasing globally, because only a few places

have been monitored carefully and even there the data are limited.

Researchers have tended to ignore or avoid jellyfish, in part because they are such a nightmare to deal with. Typical nets shred them, and collecting them intact can require heroic efforts. Shin-ichi Uye, a marine ecologist at Hiroshima University in Japan, says that because some jellyfish are heavier than sumo



Jellyfish plagued an Israeli power plant in 2011.

wrestlers, marine biologists must carefully balance their small research boats to avoid capsizing when retrieving a single specimen.

To make matters worse, many jellyfish groups have complex life cycles. Several species, including moon jellyfish, reproduce sexually to form larvae that settle on the sea floor and develop into anemone-like growths called polyps. If conditions are favourable, a single polyp can bud to form 20 floating jellies. But in hard times, polyps can produce more polyps or retreat into a tough cyst. Then, when the environment improves, the waiting polyps can fuel a massive bloom of jellyfish that might seem like an invasion from nowhere.

With very few exceptions, the polyp colonies that cause large jellyfish blooms are "really hard to find", says Claudia Mills, a jellyfish specialist at the University of Washington's Friday Harbor Laboratories, who was part of the jellyfish review group and was one of the first to examine the possibility of a global rise. Bloom triggers seem to be tied to seasonal temperature changes, she says. And that raises the possibility that warming of the oceans could indeed cause populations to mushroom⁵.

Most of the researchers who have questioned the idea of a jellyfish explosion say they cannot rule out the possibility that blooms are becoming more prevalent or that humans are at least partly responsible⁶. In Japan, for example, long-term records suggest that blooms happened only every 40 years or so before 2000, but have come nearly every year since⁷. Moreover, the blooms seem to originate in Chinese waters,

where overfishing has severely depleted the Japanese jellyfishes' main competitors.

In most other cases, however, the trends are not so clear cut. So in 2010, the biologists on the task force began a Jellyfish Database Initiative (JEDI), compiling every scientific jellyfish record they could find, and they expect to continue expanding this resource. Some researchers are also teaming up with the public. The Monterey Bay Aquarium Research Institute in California has launched a website called Jellywatch.org, through which scientists and citizens can report jellyfish sightings to help fill out the JEDI database. The intergovernmental Mediterranean Science Commission in Monaco has started a similar programme for that sea.

HIDDEN DATA

Despite the lack of long-term studies, some scientists think that there is enough evidence to answer some important questions. In a paper due out later this month, Lucas Brotz, a graduate student in fisheries biology at the University of British Columbia in Vancouver, Canada, and his adviser, Daniel Pauly, analysed media reports and other non-scientific data about bloom patterns since 1950, such as interviews with locals and scientists. The researchers used fuzzy logic, a ranking system that incorporates the reliability and abundance of information, to identify trends in less-than-ideal data sets. They found that jellyfish populations were increasing in 31 of the 45 ocean regions that they studied.

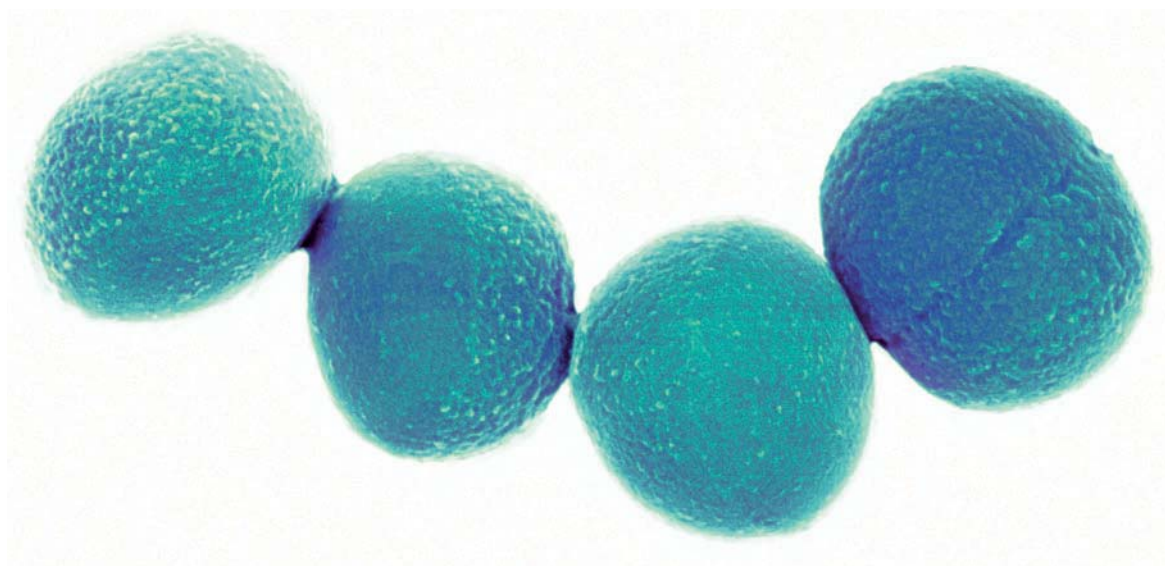
"Our study says you can actually pierce through the confusing fog of press reports and anecdotes and scientific data to establish whether increases are occurring," says Pauly.

Graham and other researchers praise the approach, but they contend that such information just isn't adequate. The jellyfish review team will begin to analyse the full JEDI database later this month, encouraged by the growth of programmes in places such as Peru and Japan, where scientists work with fishermen to monitor jellyfish populations. But even so, the researchers caution that they will need to establish many more monitoring programmes to complete a global picture.

With such programmes in place, researchers could use the comprehensive jellyfish data sets to track how the oceans are changing. These creatures make ideal environmental sentinels because humans mostly leave them alone, says Graham. "Jellyfish are the great bystanders of the oceans and the oceans' health." ■

Mark Schroppe is a freelance writer in Melbourne, Florida.

1. Condon, R. H. *et al.* *BioScience* **62**, 160–169 (2012).
2. Graham, W. M. *Hydrobiologia* **451**, 97–111 (2001).
3. Pauly, D., Christensen, V., Dalsgaard, J., Froese, R. & Torres, F. *Jr Science* **279**, 860–863 (1998).
4. Jackson, J. B. C. *Proc. Natl Acad. Sci. USA* **105**, 11458–11465 (2008).
5. Mills, C. E. *Hydrobiologia* **451**, 55–68 (2001).
6. Purcell, J. E. *Annu. Rev. Mar. Sci.* **4**, 209–235 (2012).
7. Uye, S. *Plankton Benthos Res.* **3**, S125–S131 (2008).



Man vs MRSA

For decades, Robert Daum has studied the havoc wreaked by methicillin-resistant Staphylococcus aureus. Now he thinks he can stop it for good.

BY MARYN MCKENNA

Over the years, Robert Daum has learned to respect his adversary. In 1995, he and his co-workers at the University of Chicago children's hospital in Illinois were investigating infections that had affected two dozen children in their emergency department. Three children had fast-moving pneumonia. A fourth had an abscess the size of his fist buried in the muscle of one buttock. In a fifth, the bacterium had infiltrated the bones of one foot. The infections were resistant to many common antibiotics, including methicillin. To Daum's surprise, the culprit was MRSA — methicillin-resistant *Staphylococcus aureus* — a bacterium that was thought to spread only among hospital inpatients. But none of these kids had been to the hospital for months before becoming ill.

Few researchers were willing to accept the implications. Daum wrangled for 18 months with editors at the *Journal of the American Medical Association* over a paper reporting the cases and showing that this strain was dangerous, acquired in the community and differed genetically from hospital strains. His article¹ was eventually published in 1998 and is now widely considered to be the early warning of an epidemic that currently results in millions of visits to doctors and hospitals a year².

Daum, a paediatric infectious-disease physician and founder of the University of Chicago's MRSA Research Center, is still raising the alarm about the epidemic. He sees the fight as more urgent than ever, and now thinks he knows how to win it. A few days before Christmas, he and Brad Spellberg, a physician who conducts vaccine research at the University of California, Los Angeles, published an article³ calling for a vaccine that would vanquish *S. aureus*. "We can't treat this," Daum says. "We have to prevent it."

This time, Daum's views have more support. Over the past 10 years, as MRSA has become resistant even to last-resort antibiotics, several pharmaceutical companies have launched research

programmes for vaccines, some with Daum's input. But Daum contends that they underestimate the enemy by relying on the standard immunological approach of triggering the production of protective antibodies. Instead, he advocates a strategy that stimulates T cells, part of a different branch of the immune system. It is an ambitious proposal and not all infectious-disease specialists are convinced that it will work.

"It is a provocative idea," says Gerald Pier, a microbiologist at Harvard Medical School in Boston, Massachusetts, who also works on *S. aureus* vaccines. "But it is still too early to know how applicable this component of immunity would be to vaccine development."

The MRSA bacterium (pictured) is resistant to almost all antibiotics.

EVIDENCE FOR AN EPIDEMIC

It took a while for the field to agree with Daum that a new epidemic had begun. Attitudes began to shift around 18 months after his seminal paper, when investigators for the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, reported that four children in Minnesota and North Dakota had died from infections similar to those in Daum's hospital. After that, the CDC and hospitals found clusters of community-associated infections in jails and prisons, then in sports teams, then in unusually high numbers of patients in emergency departments. Physicians, too, were reporting unexpected cases of grave illness — necrotizing pneumonia and flesh-eating disease — all caused by the community-associated strain.

From there, the epidemic grew. By one estimate, community-acquired MRSA accounts for half of the more than 14 million skin and soft-tissue infections that send people to doctors and emergency departments in the United States every year². MRSA also causes around

➔ **NATURE.COM**
Read *Nature's*
special issue on
vaccines:
go.nature.com/a3nzqx

SPL

100,000 serious blood infections and more than 15,000 deaths a year. Meanwhile, pharmaceutical manufacturers are backing away from making new antibiotics, arguing that resistance is undermining compounds too quickly for them to recoup their costs.

A study by staff at the CDC's Division of Healthcare Quality Promotion estimates that an *S. aureus* vaccine given to vulnerable groups could reduce the number of serious MRSA infections by 24,000–34,000 cases per year⁴. Researchers familiar with vaccine development say that manufacturers have recognized the potential market. “You can see it within the industry now. They are all interested,” says Jean Lee, an *S. aureus* vaccine researcher at Harvard Medical School. “Ten years ago, that wasn't the case.” But getting from interest to a viable formula is proving a formidable challenge.

SHOT IN A DIFFERENT ARM

Since Edward Jenner first scratched cowpox virus into a boy's arm 216 years ago, vaccination has mostly proceeded along variations of one strategy: introducing into the body an antigen, such as a weakened disease organism or a fragment of an organism. The immune system responds by producing an antibody, a protein that recognizes the antigen and triggers an immune attack on the organism. For years, the presence of antibodies has been taken as a sign that a person will be immune to later infection.

The first attempt at making a *S. aureus* vaccine was modelled on successful vaccines for *Streptococcus pneumoniae* and *Haemophilus influenzae*. Like them, it used antigens consisting of carbohydrate molecules from the sticky capsule surrounding the bacterium, attached to a protein produced by another bacterium. But the formula, called StaphVax and developed by Nabi Biopharmaceuticals in the late 1990s, was unsuccessful. When the company tested it in 1998–99 in a phase III trial⁵, the recipients made antibodies, but then developed *S. aureus* infections in their blood at the same rate as those who received placebo. The programme was suspended in 2005 and sold to GlaxoSmithKline in 2009. Merck, meanwhile, created a formula that used a cell-surface protein involved in the bacterium's ability to take up iron. It cancelled late-phase clinical trials in June last year because of negative results.

In retrospect, Daum says, no one should have expected an *S. aureus* vaccine to be that easy. “This organism has multiple strategies for accomplishing all its tasks, from invading the blood stream to elaborating toxins to causing local skin abscesses,” he says. “Targeting a vaccine against just one of them merely eggs the bug on.”

Unlike most pathogens, *S. aureus* is a commensal organism; it lives on the skin and in the nostrils of up to one-third of humans, mostly without causing disease.

This benign but continual occupation means that much of the population already has antibodies to the bacterium. And unlike many infections, having it once is no guarantee of protection; roughly one in four people who have one *S. aureus* infection will go on to develop another. So vaccine developers don't really know what characterizes an immune individual. “We're working without any information as to what constitutes high-level immunity to *S. aureus* infection,” says Pier.

Still, some companies are attempting to build immunity using an all-out antibody attack: Pfizer, Novartis and GlaxoSmithKline are testing formulas packed with four or five antigens, hoping to elicit an array of antibodies that will overwhelm the bacterium's defences. Several other companies are trying passive immunization — delivering antibodies harvested from people — but none has yet achieved results better than placebo.

The team that may have edged closest to success is actually on Daum's own campus. Microbiologist Olaf Schneewind, who is not affiliated with Daum's group, has developed a formula that incorporates a mutated version of a cell-wall component of *S. aureus* called protein A. Under normal circumstances, protein A binds to antibodies, protecting the bacterium from attack by the immune system; the mutant version, however, is unable to bind. So far, the formula has been tested only in mice⁶.

Daum is familiar with the failed-vaccine landscape, not just from his consulting work with pharmaceutical companies, but also from serving on the US Food and Drug Administration's Vaccine and Related Biological Products Advisory Committee. At the same time, his research into how *S. aureus* is able to cause disease in so many tissues and its bristling array of defences against the immune system convinced him that a new approach was needed.

When Daum discusses *S. aureus*, he is blunt about the challenges and impatient to move ahead. “I don't think multiple antigens are enough,” he says, sitting in the old hospital building where he spotted the first community-acquired MRSA cases, now transformed into overcrowded offices stacked with piles of journals. “I think we need multiple immunological mechanisms. And I think the central one should be something that has been considered heresy up to now.”

LUCKY BREAK

The heretical approach was inspired, in part, by a patient. As part of an ongoing project to root out the causes of recurring infections, in 2009 two of Daum's team members went to the home of a toddler who had recently been in the emergency department. But the girl wasn't there; she was in the hospital's intensive-care unit with a new infection. When Daum tracked her down, he noticed something odd in her records. She had had unusually frequent abscesses and repeated bouts of pneumonia.

Acting on a hunch, Daum teamed up with Steven Holland, chief of the clinical infectious diseases laboratory at the National Institute of Allergy and Infectious Diseases in Bethesda, Maryland, to carry out a detailed genetic analysis. Daum's hunch was right: the girl had a mutation that Holland had recently linked to a rare immunodeficiency called Job's syndrome⁷. People with the syndrome have persistent, smouldering *S. aureus* infections, owing to an inability to make a type of lymphocyte, or immune cell, called a T_H17 cell.

These cells, which make a proinflammatory protein called interleukin-17, have become a hot topic in vaccine research. They are produced by a different branch of the immune system from the one that makes antibodies, yet they still seem to be involved in the body's memory

of exposures to pathogens.

Daum believes that T_H17 cells are the key to an *S. aureus* vaccine. “It looks like T cells are very important in staphylococcal immunity,” he says. Spellberg demonstrated in 2009 that a vaccine that stimulated production of interleukin 17 could protect mice against infections of *S. aureus* and *Candida albicans*⁸. (That vaccine is now being developed by NovaDigm Therapeutics as NDV3.)

Daum and Spellberg have now joined forces and formed a cross-disciplinary team to see whether boosting the activity of T_H17 cells can prevent *S. aureus* infections in humans. The team includes an intensive-care specialist who has developed animal models to study these infections; an epidemiologist; two immunologists; and a biomedical engineer.

They have begun by selecting current and former patients with MRSA from the University of Chicago's hospitals and comparing

“This is a universal epidemic, and there should be a universal vaccine.”



Robert Daum has shown that MRSA is not confined to hospitals — and is determined to find a vaccine that will fight the bacterium.

their immunological activity with that of people who have never had MRSA. In a second phase, they will test how lymphocytes harvested from the patients react to a number of infectious organisms, including MRSA.

The project will face significant challenges, Spellberg points out. “One of the reasons the vaccine world has always been so focused on antibodies is because it’s so easy to measure antibodies,” he says. “There is no high-throughput T-cell assay. It takes a lot of work.”

FAR FROM FINAL

If the team can succeed in boosting T-cell activity, it will still be only part of a solution. The group has to consider whether to include a traditional antibody-stimulating antigen in a vaccine, and whether to add a third component, such as protein A. The researchers must also work out whether one vaccine formula can stop *S. aureus* from invading many types of tissue. “We want a vaccine that prevents invasive disease, we want a vaccine that prevents pneumonia and we want a vaccine that prevents skin infections,” Daum says. “Can one vaccine solve three separate clinical problems?”

The researchers will have to manoeuvre around *S. aureus*’s dual role as pathogen and commensal bacterium. If they wipe out the body’s benign staph occupiers, a more harmful organism might take their place.

Daum and his collaborators will also have to face down scepticism from other staph researchers, who view the T_H17 idea as intriguing but impractical. Schneewind points out that the US Food and Drug Administration’s rules require a vaccine to demonstrate antibody production to win licensure. “I am not aware of any vaccine licence

where the correlate of protective immunity is IL-17 response,” he says. But given the surge of interest in T_H17 cells, he adds, “people will try it, and we’ll see how far they get”.

Perhaps the most difficult question to answer is: who should get an *S. aureus* vaccine? Because community-acquired MRSA is so widespread, the maximum benefits might come only if the vaccine is administered to everyone. But with a great deal of suspicion of vaccines in the United States and elsewhere, an addition to the routine immunization schedule is likely to be met with resistance.

Mention these concerns to Daum, and the trademark impatience break through. “This is a universal epidemic, and there should be a universal vaccine,” he says. “I think we should put this into the paediatric vaccine schedule in the first year of life. And if it happened to work on all MRSA syndromes, like the skin infections that flood our emergency room, then we would have something wonderful on our hands.” ■

Maryn McKenna is a science journalist in Atlanta, Georgia, and author of the book *Superbug: the Fatal Menace of MRSA*.

1. Herold, B. C. *et al.* *J. Am. Med. Assoc.* **279**, 593–598 (1998).
2. Hersh, A. L. Chambers, H. F., Maselli, J. H. & Gonzales, R. *Arch. Intern. Med.* **168**, 1585–1591 (2008).
3. Daum, R. S. & Spellberg, B. *Clin. Infect. Dis.* <http://dx.doi.org/10.1093/cid/cir828> (2011).
4. Lucero, C. A. *et al.* *Vaccine* **27**, 5061–5068 (2009).
5. Shinefield, H. *et al.* *N. Engl. J. Med.* **346**, 491–496 (2002).
6. Kim, H. K., Cheng, A. G., Kim, H.-Y., Missiakas, D. M. & Schneewind, O. *J. Exp. Med.* **207**, 1863–1870 (2010).
7. Holland, S. M. *et al.* *N. Engl. J. Med.* **357**, 1608–1619 (2007).
8. Lin, L. *et al.* *PLoS Pathog.* **5**, e1000703 (2009).

COMMENT

ECOLOGY How elephants could reduce fire risk in Australia **p.30**



NEUROSCIENCE The source of the self in the brain's wiring **p.31**

LITERATURE How Charles Dickens drew on science, but left room for wonder **p.32**

OBITUARY Philip Lawley and the discovery that DNA damage can cause cancer **p.36**

ILLUSTRATION BY MARK SMITH



The toxic truth about sugar

Added sweeteners pose dangers to health that justify controlling them like alcohol, argue **Robert H. Lustig, Laura A. Schmidt and Claire D. Brindis.**

Last September, the United Nations declared that, for the first time in human history, chronic non-communicable diseases such as heart disease, cancer and diabetes pose a greater health burden worldwide than do infectious diseases, contributing to 35 million deaths annually.

This is not just a problem of the developed world. Every country that has adopted the Western diet — one dominated by low-cost, highly processed food — has witnessed rising rates of obesity and related diseases. There are now 30% more people who are obese than who are undernourished. Economic development means that the populations of low- and middle-income countries are living longer, and therefore are more

SUMMARY

- Sugar consumption is linked to a rise in non-communicable disease
- Sugar's effects on the body can be similar to those of alcohol
- Regulation could include tax, limiting sales during school hours and placing age limits on purchase

susceptible to non-communicable diseases; 80% of deaths attributable to them occur in these countries.

Many people think that obesity is the root cause of these diseases. But 20% of obese people have normal metabolism and

will have a normal lifespan. Conversely, up to 40% of normal-weight people develop the diseases that constitute the metabolic syndrome: diabetes, hypertension, lipid problems, cardiovascular disease and non-alcoholic fatty liver disease. Obesity is not the cause; rather, it is a marker for metabolic dysfunction, which is even more prevalent.

The UN announcement targets tobacco, alcohol and diet as the central risk factors in non-communicable disease. Two of these three — tobacco and alcohol — are regulated by governments to protect public health, leaving one of the primary culprits behind this worldwide health crisis unchecked. Of course, regulating food is more ►

► complicated — food is required, whereas tobacco and alcohol are non-essential consumables. The key question is: what aspects of the Western diet should be the focus of intervention?

In October 2011, Denmark chose to tax foods high in saturated fat, despite the fact that most medical professionals no longer believe that fat is the primary culprit. But now, the country is considering taxing sugar as well — a more plausible and defensible step. Indeed, rather than focusing on fat and salt — the current dietary ‘bogeymen’ of the US Department of Agriculture (USDA) and the European Food Safety Authority — we believe that attention should be turned to ‘added sugar’, defined as any sweetener containing the molecule fructose that is added to food in processing.

Over the past 50 years, consumption of sugar has tripled worldwide. In the United States, there is fierce controversy over the pervasive use of one particular added sugar — high-fructose corn syrup (HFCS). It is manufactured from corn syrup (glucose), processed to yield a roughly equal mixture of glucose and fructose. Most other developed countries eschew HFCS, relying on naturally occurring sucrose as an added sugar, which also consists of equal parts glucose and fructose.

Authorities consider sugar as ‘empty calories’ — but there is nothing empty about these calories. A growing body of scientific evidence is showing that fructose can trigger processes that lead to liver toxicity and a host of other chronic diseases¹. A little is not a problem, but a lot kills — slowly (see ‘Deadly effect’). If international bodies are truly concerned about public health, they must consider limiting fructose — and its main delivery vehicles, the added sugars HFCS and sucrose — which pose dangers to individuals and to society as a whole.

NO ORDINARY COMMODITY

In 2003, social psychologist Thomas Babor and his colleagues published a landmark book called *Alcohol: No Ordinary Commodity*, in which they established four criteria, now largely accepted by the public-health community, that justify the regulation of alcohol — unavailability (or pervasiveness throughout society), toxicity, potential for abuse and negative impact on society². Sugar meets the same criteria, and we believe that it similarly warrants some form of societal intervention.

First, consider unavailability. Evolutionarily, sugar was available to our ancestors as fruit for only a few months a year (at harvest time), or as honey, which was guarded by bees. But in recent years, sugar has been added to nearly all processed foods, limiting consumer choice³. Nature made sugar hard to get; man made it easy. In many parts of the

world, people are consuming an average of more than 500 calories per day from added sugar alone (see ‘The global sugar glut’).

Now, let’s consider toxicity. A growing body of epidemiological and mechanistic evidence argues that excessive sugar consumption affects human health beyond simply adding calories⁴. Importantly, sugar induces all of the diseases associated with metabolic syndrome^{1,5}. This includes: hypertension (fructose increases uric acid, which raises blood pressure); high triglycerides and insulin resistance through synthesis of fat in the liver; diabetes from increased liver

“Sugar is cheap, sugar tastes good and sugar sells, so companies have little incentive to change.”

glucose production combined with insulin resistance; and the ageing process, caused by damage to lipids, proteins and DNA through non-enzymatic binding of fructose to these molecules. It can also be argued that fructose exerts toxic effects on the liver that are similar to those of alcohol¹. This is no surprise, because alcohol is derived from the fermentation of sugar. Some early studies have also linked sugar consumption to human cancer and cognitive decline.

Sugar also has clear potential for abuse. Like tobacco and alcohol, it acts on the brain to encourage subsequent intake. There are now numerous studies examining the dependence-producing properties of sugar in humans⁶. Specifically, sugar dampens the suppression of the hormone ghrelin, which signals hunger to the brain. It also interferes with the normal transport and signalling of the hormone leptin, which helps to produce the feeling of satiety. And it reduces dopamine signalling in the brain’s reward centre, thereby decreasing the pleasure derived from food and compelling

the individual to consume more^{1,6}.

Finally, consider the negative effects of sugar on society. Passive smoking and drink-driving fatalities provided strong arguments for tobacco and alcohol control, respectively. The long-term economic, health-care and human costs of metabolic syndrome place sugar overconsumption in the same category⁷. The United States spends \$65 billion in lost productivity and \$150 billion on health-care resources annually for morbidities associated with metabolic syndrome. Seventy-five per cent of all US health-care dollars are now spent on treating these diseases and their resultant disabilities. Because about 25% of military applicants are now rejected for obesity-related reasons, the past three US surgeons general and the chairman of the US Joint Chiefs of Staff have declared obesity a “threat to national security”.

HOW TO INTERVENE

How can we reduce sugar consumption? After all, sugar is natural. Sugar is a nutrient. Sugar is pleasure. So too is alcohol, but in both cases, too much of a good thing is toxic. It may be helpful to look to the many generations of international experience with alcohol and tobacco to find models that work^{8,9}. So far, evidence shows that individually focused approaches, such as school-based interventions that teach children about diet and exercise, demonstrate little efficacy. Conversely, for both alcohol and tobacco, there is robust evidence that gentle ‘supply side’ control strategies which stop far short of all-out prohibition — taxation, distribution controls, age limits — lower both consumption of the product and the accompanying health harms. Successful interventions share a common end-point: curbing availability^{2,8,9}.

Taxing alcohol and tobacco products — in the form of special excise duties, value-added taxes and sales taxes — are the most popular and effective ways to

DEADLY EFFECT

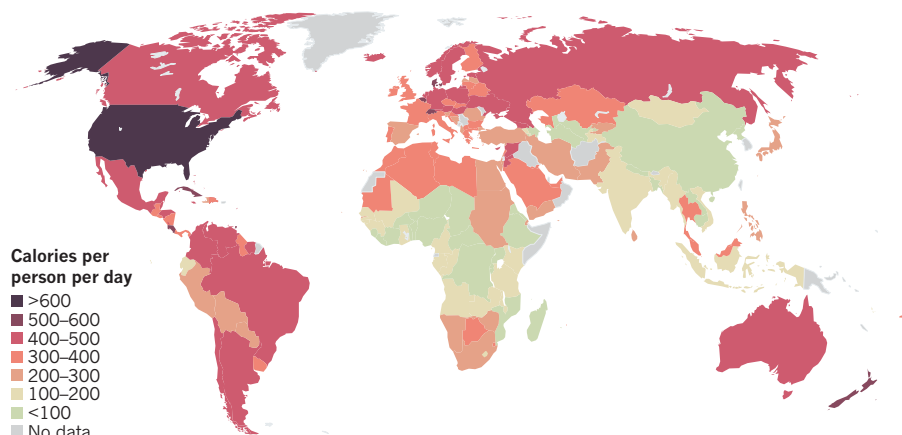
Excessive consumption of fructose can cause many of the same health problems as alcohol.

Chronic ethanol exposure	Chronic fructose exposure
Haematological disorders	
Electrolyte abnormalities	
Hypertension	Hypertension (uric acid)
Cardiac dilatation	
Cardiomyopathy	Myocardial infarction (dyslipidaemia, insulin resistance)
Dyslipidaemia	Dyslipidaemia (<i>de novo</i> lipogenesis)
Pancreatitis	Pancreatitis (hypertriglyceridaemia)
Obesity (insulin resistance)	Obesity (insulin resistance)
Malnutrition	Malnutrition (obesity)
Hepatic dysfunction (alcoholic steatohepatitis)	Hepatic dysfunction (non-alcoholic steatohepatitis)
Fetal alcohol syndrome	
Addiction	Habituation, if not addiction

Source: ref. 1

THE GLOBAL SUGAR GLUT

Global sugar supply (in the form of sugar and sugar crops, excluding fruit and wine) expressed as calories per person per day, for the year 2007.



reduce smoking and drinking, and in turn, substance abuse and related harms². Consequently, we propose adding taxes to processed foods that contain any form of added sugars. This would include sweetened fizzy drinks (soda), other sugar-sweetened beverages (for example, juice, sports drinks and chocolate milk) and sugared cereal. Already, Canada and some European countries impose small additional taxes on some sweetened foods. The United States is currently considering a penny-per-ounce soda tax (about 34 cents per litre), which would raise the price of a can by 10–12 cents. Currently, a US citizen consumes an average of 216 litres of soda per year, of which 58% contains sugar. Taxing at a penny an ounce could provide annual revenue in excess of \$45 per capita (roughly \$14 billion per year); however, this would be unlikely to reduce total consumption. Statistical modelling suggests that the price would have to double to significantly reduce soda consumption — so a \$1 can should cost \$2 (ref. 10).

Other successful tobacco- and alcohol-control strategies limit availability, such as reducing the hours that retailers are open, controlling the location and density of retail markets and limiting who can legally purchase the products^{2,9}. A reasonable parallel for sugar would tighten licensing requirements on vending machines and snack bars that sell sugary products in schools and workplaces. Many schools have removed unhealthy fizzy drinks and candy from vending machines, but often replaced them with juice and sports drinks, which also contain added sugar. States could apply zoning ordinances to control the number of fast-food outlets and convenience stores in low-income communities, and especially around schools, while providing incentives for the establishment of grocery stores and farmer's markets.

Another option would be to limit sales during school operation, or to designate an age limit (such as 17) for the purchase of drinks with added sugar, particularly soda. Indeed, parents in South Philadelphia, Pennsylvania, recently took this upon themselves by lining up outside convenience stores and blocking children from entering them after school. Why couldn't a public-health directive do the same?

THE POSSIBLE DREAM

Government-imposed regulations on the marketing of alcohol to young people have been quite effective, but there is no such approach to sugar-laden products. Even so, the city of San Francisco, California, recently banned the inclusion of toys with unhealthy meals such as some types of fast food. A limit — or, ideally, ban — on television commercials for products with added sugars could further protect children's health.

Reduced fructose consumption could also be fostered through changes in subsidization. Promotion of healthy foods in US low-income programmes, such as the Special Supplemental Nutrition Program for Women, Infants and Children and the Supplemental Nutrition Assistance Program (also known as the food-stamps programme) is an obvious place to start. Unfortunately, the petition by New York City to remove soft drinks from the food-stamp programme was denied by the USDA.

Ultimately, food producers and distributors must reduce the amount of sugar added to foods. But sugar is cheap, sugar tastes good and sugar sells, so companies have little incentive to change. Although one institution alone can't turn this juggernaut around, the US Food and Drug Administration could "set the table" for change⁸. To start, it should consider removing fructose from the Generally Regarded as Safe

(GRAS) list, which allows food manufacturers to add unlimited amounts to any food. Opponents will argue that other nutrients on the GRAS list, such as iron and vitamins A and D, can also be toxic when over-consumed. However, unlike sugar, these substances have no abuse potential. Removal from the GRAS list would send a powerful signal to the European Food Safety Authority and the rest of the world.

Regulating sugar will not be easy — particularly in the 'emerging markets' of developing countries where soft drinks are often cheaper than potable water or milk. We recognize that societal intervention to reduce the supply and demand for sugar faces an uphill political battle against a powerful sugar lobby, and will require active engagement from all stakeholders. Still, the food industry knows that it has a problem — even vigorous lobbying by fast-food companies couldn't defeat the toy ban in San Francisco. With enough clamour for change, tectonic shifts in policy become possible. Take, for instance, bans on smoking in public places and the use of designated drivers, not to mention airbags in cars and condom dispensers in public bathrooms. These simple measures — which have all been on the battleground of American politics — are now taken for granted as essential tools for our public health and well-being. It's time to turn our attention to sugar. ■

Robert H. Lustig is in the Department of Pediatrics and the Center for Obesity Assessment, Study and Treatment at the University of California, San Francisco, California 94143, USA. **Laura A. Schmidt** and **Claire D. Brindis** are at the Clinical and Translational Science Institute and the Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, California 94118, USA. e-mail: rlustig@peds.ucsf.edu

1. Lustig, R. H. *J. Am. Diet. Assoc.* **110**, 1307–1321 (2010).
2. Babor, T. et al. *Alcohol: No Ordinary Commodity: Research and Public Policy* (Oxford Univ. Press, 2003).
3. Vio, F. & Uauy, R. in *Food Policy for Developing Countries: Case Studies* (eds Pinstrup-Andersen, P. & Cheng, F.) No. 9-5 (2007); available at <http://go.nature.com/prjsk4>
4. Joint WHO/FAO Expert Consultation. *Diet, Nutrition and the Prevention of Chronic Diseases* WHO Technical Report Series 916 (WHO, 2003).
5. Tappy, L., Lê, K. A., Tran, C. & Paquot, N. *Nutrition* **26**, 1044–1049 (2010).
6. Garber, A. K. & Lustig, R. H. *Curr. Drug Abuse Rev.* **4**, 146–162 (2011).
7. Finkelstein, E. A., Fiebelkorn, I. C. & Wang, G. *Health Aff.* **W3** (suppl.), 219–226 (2003).
8. Engelhard, C. L., Garson, A. Jr & Dorn, S. *Reducing Obesity: Policy Strategies from the Tobacco Wars* (Urban Institute, 2009); available at <http://go.nature.com/w4o5uk>
9. Room, R., Schmidt, L. A., Rehm, J. & Mäkelä, P. *Br. Med. J.* **337**, a2364 (2008).
10. Sturm, R., Powell, L. M., Chiqui, J. F. & Chaloupka, F. J. *Health Aff.* **29**, 1052–1058 (2010).

Bring elephants to Australia?

There's a solution to the continent's rampant fires and feral animals, says **David Bowman** — introduce large mammals and increase hunting pressure.

Three years ago this week, Australia was burning. On 7 February 2009 — now known as Black Saturday — a massive firestorm consumed more than 400,000 hectares in southern Australia. At least 173 people died trying to outrun the fires, defend their homes or seek shelter.

That blaze was unusually fierce, but fires are a constant source of anxiety for Australia. The continent is extremely fire-prone, with a distinctive signature of oscillating fire activity that begins in the north during the winter, then moves south during the summer. Lately, the fires have been more intense and widespread, perhaps as a result of climate change — last year, around 5% of the continent was burnt.

If only fires were Australia's sole environmental concern. The continent is also overrun by invasive species. They fill holes created by a mass extinction event that occurred around 50,000 years ago during the Pleistocene, when the arrival of the first Australians coincided with a collapse in the continent's megafauna, namely giant marsupials (some as large as hippopotamuses), reptiles and birds¹. The precise causes of that event are unclear, but the resulting gap in the food web has been filled by populations of pigs, goats, cattle, horses, donkeys, camels, buffalo and deer². These animals are reconstructing ecosystems, a trend amplified by the introduction of alien plants, particularly ultra-competitive grasses³.

So far, Australia and other regions dealing with similar environmental crises have responded by spending millions on piecemeal control programmes that target one problem at a time. For instance, Australians have tried attaching radio collars to buffalo and then following them in helicopters to find and kill others in the herd³. Such approaches cost millions, and are ultimately ineffectual — for every buffalo hunted, another is born to replace it.

I think that another, more holistic approach can address Australia's ecological problems. Specifically, we must restabilize food webs (now out of balance because of the Pleistocene extinctions), the loss of the Aboriginal traditions of patch burning and hunting, and the ad hoc release of non-native animals and plants. We must introduce and manage predators to control the feral animals, and bring in herbivore species to graze the flammable grasses — which we can better control using small fires as 'über-herbivores'⁴.

I accept that this is a radical way of thinking; we would have to weigh the various options. For example, we could stop poisoning the Australian wolf (dingo). Poisoning disrupts their social structure, and research suggests that dingoes in packs act as top predators of smaller predators such as introduced foxes⁵. More dingoes could also help to control other feral animals, such as



pigs. Alternatively, we could introduce predators such as the Komodo dragon, which would fill the niche of the giant lizards that once thrived in Australia⁶.

But, because such predators can pose threats to humans and livestock, it might be more realistic to employ Aboriginal hunters, who could help to control feral animals and restore the traditional practice of patch burning. Indeed, existing ranger programmes that enable indigenous people to return to their roots — by hunting buffalo or managing natural resources — have been shown to have social and health benefits for this disadvantaged sector of the Australian community⁷.

A major source of fuel for wildfires in the monsoon tropics is gamba grass, a giant African grass that has invaded north Australia's savannas⁸. It is too big for marsupial grazers (kangaroos) and for cattle and buffalo, the largest feral mammals. But gamba grass is a great meal for elephants or rhinoceroses. The idea of introducing elephants may seem absurd, but the only other methods likely to

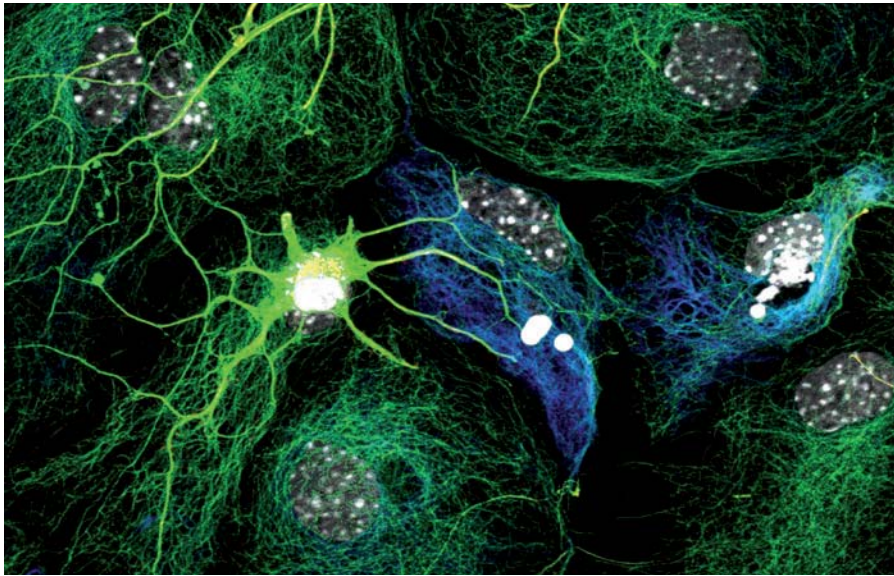
control gamba grass involve using chemicals or physically clearing the land, which would destroy the habitat. Using mega-herbivores may ultimately be more practical and cost-effective, and it would help to conserve animals that are threatened by poaching in their native environments. This potential solution is not limited to Australia — it has been suggested that elephants could be used as part of a project to 'rewild', or return North American ecosystems to their prehuman state⁹.

I realize that there are major risks associated with what I am proposing. It would be essential to proceed cautiously, with well-designed studies to monitor the effects. The greatest challenge would be managing the density of herbivore populations so that their demand on resources does not degrade the ecosystem. Here, we could adopt management methods from game parks and reserves, such as building fences, regulating the availability of water and food, and controlling breeding and hunting.

Of course, introducing large mammals cannot solve all of Australia's ecological-management conundrums. And I am mindful that the proposal could be used to justify commercial grazing in fragile ecosystems, an ongoing controversy¹⁰. But the usual approaches to managing these issues aren't working. The full spectrum of options needs to be canvassed in an open and honest way. ■

David Bowman is professor of environmental change biology at the School for Plant Science, University of Tasmania, Private Bag 55 Hobart, Tasmania 7001, Australia.
e-mail: david.bowman@utas.edu.au

1. Johnson, C. *Australia's Mammal Extinctions: A 50,000-Year History* (Cambridge Univ. Press, 2006).
2. Low, T. *Feral Future: The Untold Story of Australia's Exotic Invaders* (Univ. Chicago Press, 1999).
3. Finlayson, C. M., Storrs, M. J. & Lindner, G. *Wetlands Ecol. Manage.* **5**, 19–36 (1997).
4. Bond, W. J. & Keeley, J. E. *Trends Ecol. Evol.* **20**, 387–394 (2005).
5. Johnson, C. N., Isaac, J. L. & Fisher, D. O. *Proc. R. Soc. B* **274**, 341–346 (2007).
6. Flannery, T. *The Future Eaters: An Ecological History of the Australasian Lands and People* (Reed Books, 1994).
7. Burgess, C. P. et al. *Med. J. Australia* **190**, 567–572 (2009).
8. Setterfield, S. A. et al. *Divers. Distrib.* **16**, 854–861 (2010).
9. Donlan, J. et al. *Nature* **436**, 913–914 (2005).
10. Williams, R. J., Wahren, C.-H., Bradstock, R. A. & Müller, W. J. *Austral. Ecol.* **31**, 925–936 (2006).



A microscope image of brain glia and neurons reveals the tangle of connections in this mental 'jungle'.

NEUROSCIENCE

The connected self

Christof Koch tours the idea that the essence of the mind lies in the links between neurons.

We moderns believe that our momentary, subjective experience is intimately linked to events in the brain. One set of neurons fires, and we perceive an apple's colour, while a different population of cells gives rise to its taste. Yet the self is also stable: turn the brain off, as happens during heart surgery when the body is cooled to frigid temperatures, and on recovery, the patient's character, personality, habits and long-term memories remain intact. It is these stable aspects of the self, rather than the ebb and flow of our thoughts and percepts, that physicist-turned-neuroscientist Sebastian Seung seeks to explain in *Connectome*.

Seung argues intelligently and powerfully that the self lies in the totality of the brain's wiring — the eponymous 'connectome', the word used by neuroscientists to denote all the fibre bundles (the white matter) of the human brain. These insulated nerve axons have a total length of around 150,000 kilometres. Seung hails a new science, 'connectomics', as the key to understanding the brain and its pathologies.

This view is grounded in a older doctrine known as connectionism, which postulates that neurons are simple devices and that their connections determine their functions. Cataloguing the links among neurons therefore charts the mind.

► **NATURE.COM**
For *Nature's* special issue on autism:
go.nature.com/ypo9f8

The heart of *Connectome* deals with how nervous systems can be reconstructed using electron microscopy. To do this, neural tissue is cut into slices 40–50 nanometres thick, and then imaged to a resolution of a few nanometres. Imaging 1 cubic millimetre of cortex generates 1 petabyte of data, or about a billion photo images from a typical digital camera.

Yet taking the pictures is not the hardest part. Far more daunting is labelling and tracing each neuronal process as it wends its way through the stack of images. Done manually, this would take a trained technician a million working years for each cubic millimetre of brain; luckily, computer vision and machine-learning algorithms speed things up. Last year, two teams reported in *Nature* partial reconstructions of small volumes in the mouse retina and cortex.

Seung's companion through the neural labyrinth is Hollywood star Jennifer Aniston. In 2005, my group, collaborating with neurosurgeon Itzhak Fried, discovered individual neurons that respond selectively to the image



Connectome: How the Brain's Wiring Makes Us Who We Are
SEBASTIAN SEUNG
Houghton Mifflin
Harcourt: 2012.
384 pp. \$27

or the name of familiar or famous individuals, such as your grandmother or Aniston. Seung reasons that such neurons must receive input from many less specific cells that respond to eye and hair colour and so on. These in turn connect to neurons responding to even more elementary stimuli, such as vertical lines. The essence of a person's concept of a celebrity lies, then, not with the individual neuron, but in its wiring. And this, in Seung's view, is why reconstructing connections is the key to understanding the mind.

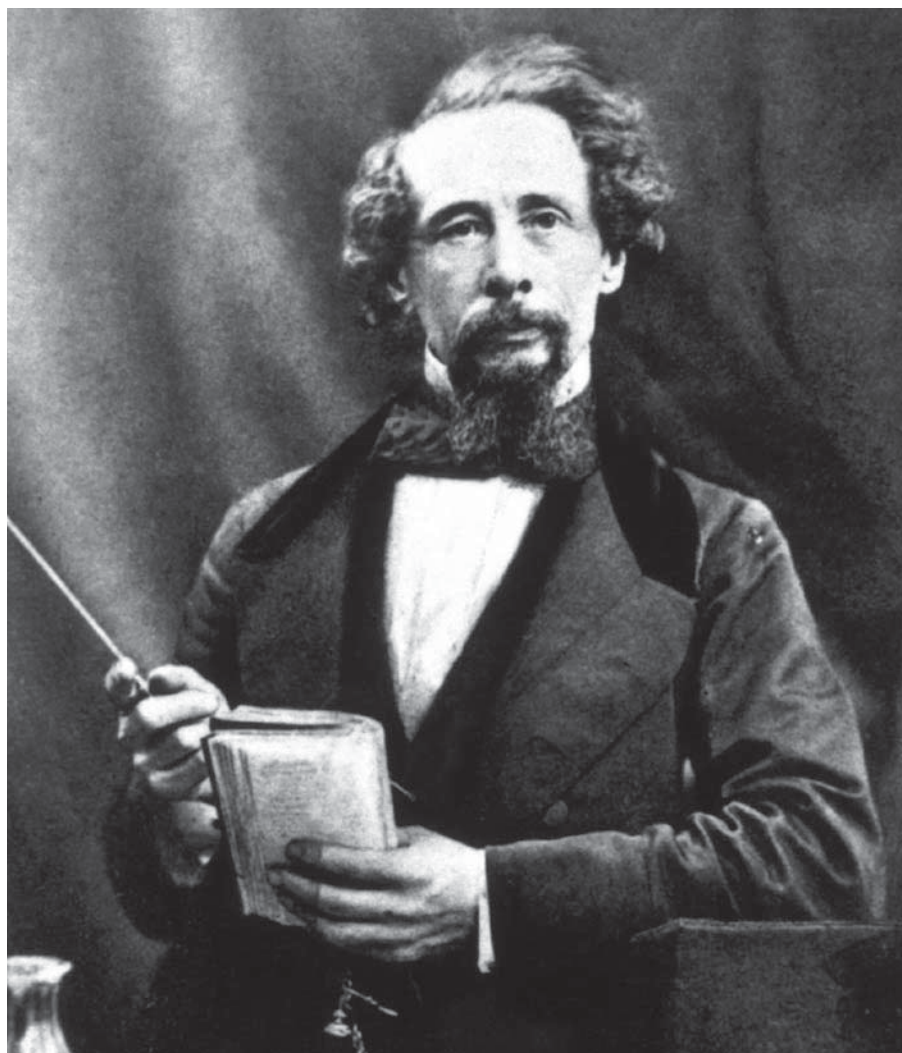
Not even Aniston's agent would think that people are born with neurons that recognize her. Rather, the neural networks that feed into these highly selective cells have wired themselves up to respond in this way. Seung discusses what we know about such plasticity, describing how mammalian nervous systems rewire themselves to adapt to the environment, balance the competing demands of plasticity and stability (so new memories don't displace old ones) and compensate for damage by rerouting information.

Treating the connectome as the be-all and end-all of brain function has its problems. Seung, for example, rebrands autism and schizophrenia as 'connectopathies' — diseases in which the brain's wiring goes awry. Yet plenty of other things are wrong in brains with these disorders besides their connectivity.

Faults in synaptic transmission and in processes inside neurons and the glial cells that support them have all been implicated in mental illness and brain disease. Neurons are intricate devices with elaborate input structures that show complex, time-dependent and nonlinear processing. They have various characteristic, and often tortuous, morphologies. Connectionism treats all this as irrelevant. Even though we have known the connectome of the nematode worm for 25 years, we are far from reading its mind. We don't yet understand how its nerve cells work.

The book is well illustrated and sourced, with an ending that is both engaging and idiosyncratic. The last two chapters are predicated on the idea that "you are your connectome", as Seung quips. One deals with technologies to freeze and pickle dead brains long term, in the hope of one day reviving their minds. The final chapter looks at the possibility of simulating the entire human brain on a future supercomputer, allowing you to upload yourself and live on as a software construct. This is the dream of the singularity that Ray Kurzweil and other technologists are hoping will render them immortal: the rapture for nerds. ■

Christof Koch is chief scientific officer at the Allen Institute for Brain Science, Seattle, Washington, USA, and a professor of biology and engineering at the California Institute of Technology, Pasadena, California, USA.
e-mail: koch@klab.caltech.edu



Novelist Charles Dickens delighted in scientific revelations, but deplored sterile reductionism.

LITERATURE

Wonders and ologies

In the week of the Dickens bicentennial, **Alice Jenkins** explores the literary giant's conflicted take on science.

Go and be somethingological directly," Mrs Gradgrind orders her children in Charles Dickens's *Hard Times* (1854). The children are brought up on their father's strict utilitarian principles — or rather, on Dickens's ferocious satire of them. So they have "a little conchological cabinet, and a little metallurgical cabinet, and a little mineralogical cabinet; and all the specimens were all arranged and labelled", but they have no fairy tales

or nursery rhymes — nothing that might encourage fancy or wonder.

Not surprisingly, the children come to bad ends. On her deathbed, Mrs Gradgrind tells her daughter: "There is something — not an Ology at all — that your father has missed, or forgotten."

Science, in Dickens's view, does immense good — moral, social and intellectual — but only when it works hand in hand with imagination and reverence. Relations between

science and Christianity in the nineteenth century were often more harmonious than we might imagine, if we focus only on the challenges that natural selection posed to some kinds of religious belief. Dickens is an interesting case study.

He was a Protestant Christian, but had no strong affiliation to any particular sect, and did not see science as a threat to religious faith. On the contrary, he argued, learning the true nature of forces or objects brings us closer to their creator. In a speech he gave in 1869 at the Birmingham and Midland Institute, he speculated that Jesus might have taught scientific truths about the "wonders on every hand", but chose not to because "the people of that time could not bear them".

It is characteristic of Dickens's undogmatic attitude to both science and religion that he was largely unfazed by Charles Darwin's *On the Origin of Species* (John Murray, 1859). He published a review of the *Origin* in his magazine, *All the Year Round*, in 1860. Although not wholly persuaded by Darwin, the author did acknowledge that the theory "entails the vastest consequences", and quoted Darwin at length. Darwin in turn was a long-time fan of Dickens's novels, and literary critic Gillian Beer has suggested that Darwin drew on Dickens in writing the *Origin*. In *Darwin's Plots* (Cambridge University Press; 2000) Beer highlights the shared concerns of these two eminent men — among them, the relationship between the extraordinary profusion of people and things, and the many-layered interconnections between entities.

SCIENCE WITH FEELING

Dickens's objection in *Hard Times* was not to science itself, but to the reductionist principle that imposes stultifying order and leaves no room for emotion or imagination. Plenty of Victorian scientific writers would have agreed with him. Michael Faraday, for example, taught that "in the pursuit of physical science, the imagination should be taught to present the subject investigated in all possible, and even in impossible views". And, in a passage that became a touchstone for Victorian science writing, geologist Adam Sedgwick wrote in 1831 that if geology were to cause "the imagination, the feelings" to be "blunted and impaired", then the subject would become "little better than a moral sepulchre".

Dickens was appalled by people whose scientific knowledge was not connected to imagination or feelings. As soon as we meet Bradley Headstone, the teacher in *Our Mutual Friend* (1865), we know that he will prove a villain, because his mind is

➔ **NATURE.COM**
Gillian Beer's essay
on time in Lewis
Carroll's *Alice* books:
go.nature.com/p8kq1d

CHARLES DICKENS MUSEUM, LONDON

rule-bound and sterile: “From his earliest childhood up, his mind had been a place of mechanical stowage ... astronomy to the right, political economy to the left — natural history, the physical sciences, figures, music, the lower mathematics, and what not, all in their several places.” The same tidy-mindedness that indicates the barrenness of the little Gradgrinds’ natural specimens foretells Headstone’s descent into criminal insanity.

What excited Dickens most about science was its ability to reveal an unimaginable world behind ordinary objects. “The facts of science are at least as full of poetry, as the most poetical fancies,” he wrote in an 1848 review of Robert Hunt’s *The Poetry of Science*. By revealing the wonder of everyday things, science compensates us for the beloved but ignorant beliefs it destroys. “When [science] has freed us from a harmless superstition,” Dickens wrote in the same review, “she offers to our contemplation something better and more beautiful, something which, rightly considered, is more elevating to the soul, nobler and more stimulating to the soaring fancy.” Dinosaurs, he went on, are really far more impressive than dragons, and coral reefs more so than mermaids.

Accordingly, Dickens championed writers who used science to show the world as spectacular, magical or astonishing. Among his close friends were mathematician and father of the computer Charles Babbage, and Richard Owen, the comparative anatomist who coined the word dinosaur. Owen was enthralled by Dickens’s novels, following them avidly as they came out in instalments. Science historian Gowan Dawson has argued that there are strong similarities between Owen’s palaeontological studies

MARY EVANS PICTURE LIBRARY



Palaeontology pioneer Richard Owen (with a moa skeleton) was one of Dickens’s close friends.



Dickens wove science into *Our Mutual Friend*, basing taxidermist Mr Venus on anatomist Richard Owen.

— which followed the principles of French naturalist Georges Cuvier to reconstruct entire extinct animals from a single fossil bone — and the way Victorian novel readers tackled serially published novels.

Both Owen and Babbage seem to have made their way into Dickens’s novels — Owen as the taxidermist Mr Venus in *Our Mutual Friend* (1865), and Babbage as one of the models for Daniel Doyce, the gifted inventor in *Little Dorrit* (1857). Dickens uses his satire of the ‘Circumlocution Office’ in this novel to protest against the sluggish government bureaucracy that was delaying Babbage’s progress with his Difference Engine.

HOUSEHOLD NAMES

Scientists also appeared in Dickens’s journalism. As editor of the popular family magazine *Household Words* (1850–59), Dickens sought out writers who he thought were sympathetic to his view that science should make us see our everyday surroundings imaginatively. In 1850, for instance, he asked Percival Leigh, a physician-turned-writer, to rework Faraday’s notes for his lectures on the candle and domestic chemical philosophy as articles for the magazine.

Leigh turned Faraday’s notes into narratives, inventing a family, the Wilkinsons, whose young son regaled his family at tea-time with information he had acquired from the scientist’s lectures at the Royal Institution. This format meant that the science could be broken up into conversational chunks, and lightened with mild domestic comedy.

Even though Dickens was happy to endorse contemporary science when he judged it to be supporting religion, feeding the imagination and telling stories, he was not above flouting scientific law for the sake of sensation. In *Bleak House* (1853), for example, two men in search of a crucial lost bundle of letters visit

Krook’s rag-and-bottle shop, only to find that “a smouldering suffocating vapour”, “a dark greasy coating on the walls and ceiling”, and a thing that looks like a small burnt log are all that remain of Krook: he has been the victim of “Spontaneous Combustion”.

The controversy that followed the publication of this unscientific episode is well known. Reproached in print by science writer George Henry Lewes for perpetuating a “vulgar error ... peculiarly adapted to the avid credulity of unscientific minds”, Dickens responded with a list of apparent real-world cases of spontaneous combustion and a defiant preface defending them as authentic, even though they had been thoroughly discredited by Lewes and others.

Aware that this was not enough to regain credibility, Dickens concluded his preface with an appeal to the imagination: “In *Bleak House*, I have purposely dwelt upon the romantic side of familiar things.”

It was the same argument that he had always made: that everyday things, and ordinary people, contain the potential for astonishing transformation. In the past he had championed science as a way of revealing this “romantic side”, but this time, backed into a corner, he used it to defend a belief that no man of science could countenance.

For Dickens, science was compelling when it could be domesticated, moralized and made into an updated version of the old fairy tales, a way of telling poetical and magical tales about the world. But when science conflicted with a good story — he combusted it. ■

Alice Jenkins is a professor of Victorian literature and culture at the University of Glasgow, UK, and author of *Space and the ‘March of Mind’: Literature and the Physical Sciences in Britain, 1815–1850*. e-mail: alice.jenkins@glasgow.ac.uk



Q&A Gary Marcus

The instrumentalist

Cognitive psychologist Gary Marcus researches language acquisition and the evolution of the brain at New York University. On the release of his latest book, *Guitar Zero — a personal yet scientific investigation of how people become musical* — he explains how a love of music and a computer game helped him to overcome a lack of rhythm and learn to play the guitar.

What were your early experiences of music?

As a child, I got kicked out of music class when I couldn't play *Mary Had a Little Lamb* on my recorder. Years later, Daniel Levitin, author of *This is Your Brain on Music* [Atlantic, 2007], tried to teach me the guitar, but discovered that I was so rhythmically impaired that I couldn't even follow a metronome. Then I discovered the computer game *Guitar Hero*. I was terrible at first, but luckily computers are patient. I got better at it, and for the first time felt the barest inkling of rhythm. Coupled with a love of listening to music, I was inspired to try playing the guitar for real.

How did you progress?

I taught myself to begin with, playing for an hour or two every day. After learning simple chords, I discovered the pentatonic scale, a set of notes that sound good in any order. I read every guitar book I could, worked with ear-training apps and experimented with music-creation software to get a sense of composition and arrangement. I tried to make up for limitations in my ears and fingers by developing a basic understanding of music theory. After six months I began to study with a teacher. Observant teachers target your weakest skills, as Swedish psychologist Anders

Guitar Zero: The New Musician and the Science of Learning

GARY MARCUS
Penguin Press: 2012.
274 pp. \$25.95

I practised a lot with a drum machine, and wrote an iPhone app called Chatternome that counts each beat out loud. That really helped me to internalize a sense of rhythm.

What are your musical goals?

To discover the underlying logic of music, to improvise rather than memorize. Our brains are set up to like new information; the reward I get from learning chord progressions and techniques has kept me going.

Why is the guitar so hard to learn?

It is partly because of the demands it places on memory. Because each pitch appears in multiple places on the fretboard, and our brains don't have the kind of database memory computers do, novices are easily confused. Playing each note requires coordination between the left and right hands. One hand has to be agile and strong to play chords, the other precise enough to play complex fingerpicking patterns. You

Ericsson recommends. Incremental learning is key, especially for adults.

How did you improve your sense of rhythm?

also have to overcome the brain's natural tendency to get sloppy as you go faster.

Does your academic research involve music?

With neurobiologist Ofer Tchernichovski of Hunter College in New York, I'm researching how songbirds learn musical syntax, and comparing that to how children learn language. I am also looking at how people understand music in time, and how children learn the implicit underpinnings of music theory — the latter using a game I've written for the iPad, and hope to release for free later in 2012. And, with friends, I am building a nylon-string guitar that can safely be played in a magnetic resonance imaging (MRI) machine. We know something about how people's brains function when playing the piano, but much less about what happens when someone plays the guitar.

Name a favourite music experiment.

Psychologists Bruno Repp at Haskins Laboratories in New Haven, Connecticut, and Peter Keller at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany, asked skilled pianists to sight-read classical compositions on completely muted electronic keyboards. Later, the pianists were able to pick out their own renditions from audio playbacks of a number of performances. This highlights the exquisite way in which experts align the neural representations of their fingers, ears and brain.

Is music the product of natural selection?

I am doubtful. Proponents like to give anecdotes about how famous musicians have more children, but rarely mention all the oboe teachers who dine alone. When you look at the brain, there is no 'music module' as there is for face recognition. Music is spread throughout the brain and implemented by many regions, such as Broca's area and the amygdala, that mainly serve other functions such as language and emotion. Music is a kind of technology that has developed over time, rather than something that is built into our minds.

What's more important, talent or practice?

Neither. If you look at the data, only about half of the variance in musical skill is accounted for by the time spent practising. Some people just learn faster and reach higher peaks. Many genes have indirect effects on music, including genes that influence auditory processing, extroversion and dexterity. All of these contribute to talent. To be truly great, you need both talent and practice. To be merely good — like me — you can get away with just practice. ■

INTERVIEW BY JASCHA HOFFMAN

Correspondence

Asian medicine: Japan's paradigm

The international medical community could benefit from the wide range of therapeutic options that traditional Japanese Kampo medicine can offer. Its integration into modern medicine has already been realized in Japan (*Nature* **480**, S96; 2011), where it is available as a 5-year specialization for physicians already trained in Western medicine.

Kampo and traditional Chinese medicine have common roots, but Kampo uses additional diagnostic techniques and more rigorously controls the quality of herbal preparations.

It would be a major loss for both Western and traditional medicines if political or financial factors were to cause the "sun to set" on Kampo.

Silke Cameron *University Medicine Göttingen, Germany.*
silke.cameron@med.uni-goettingen.de

Heidrun Reissenweber *Clinic for Japanese Medicine, Munich, Germany.*

Kenji Watanabe *Center for Kampo Medicine, Keio University School of Medicine, Tokyo, Japan.*

Asian medicine: a fungus in decline

Estimates of wildlife trade for traditional Asian medicine should include that of the caterpillar fungus *Ophiocordyceps sinensis* (*Nature* **480**, S101–S103; 2011).

The fungus, used to treat asthma and other diseases, is legally harvested on a huge scale in Tibet and the Himalayas, and is one of the world's most expensive natural medical resources. Some 85–185 tonnes are collected annually by the local population for a global market worth between US\$5 billion and \$11 billion.

Large increases in the price (up by 900% from 1997 to 2008)

and trade of caterpillar fungus have encouraged more intensive harvesting. My informal survey of harvesters in the Himalayas reveals that caterpillar fungus abundance is dwindling: the average harvest per collector dropped by around half between 2006 and 2010. Harvesters are extending their range as a result, risking overexploitation of a pristine landscape and more ecosystem degradation.

Conservation efforts must be initiated to halt the decline of this species, which is causing a loss of biodiversity and threatening local livelihoods.

Uttam Babu Shrestha *University of Massachusetts, Boston, USA.*
ubshrestha@yahoo.com

Asian medicine: many unique types

The different branches of traditional Asian medicine are frequently confused (*Nature* **480**, S81–S103; 2011). Now could be the time to revive the centuries-old term 'Eastern medicine' to avoid such inaccuracies and to complement descriptions of Western medicine.

'Oriental' and 'Asian' medicine collectively describe the range of traditional treatments used in many Asian countries. Traditional Chinese medicine is more specific. Although practised mainly in China, it influenced the development of traditional medicines unique to Japan, Korea and Vietnam in the past few hundred years. Lumping all of these together as 'traditional Chinese medicine' is therefore incorrect.

The term 'Eastern medicine' was first coined in 1613 by a court physician in Korea, Heo Jun, in his book *Donguibogam* ('Principles and Practice of Eastern Medicine'). The book is still used in clinics and, in 2009, was added to the United Nations Educational, Scientific and Cultural Organization's World Documentary Heritage list.

Hoyun Lee *Health Sciences North, Sudbury, Ontario, Canada.*
mhlee@hsnsudbury.ca

Asian medicine: call for more safety data

Marketing of traditional Chinese medicines is developing rapidly worldwide (*Nature* **480**, S81–S103; 2011). So much so that the European Union (EU) issued a directive in 2004 that all herbal preparations should be subject to the same screening procedures as pharmaceuticals by 2011. But by April last year, no Chinese herbal medicines had met the directive's requirements. Many have therefore been withheld from sale in the EU.

If these traditional remedies are to be accepted, their quality needs to be standardized and rigorous scientific data must be supplied on their efficacy and safety. The mystique surrounding such treatments must give way to verification and a proper understanding of concepts and applications. Only then can traditional Chinese medicine be integrated into a global health-care system.

Juncai Xu, Min Liu *Shanghai Clinical Research Center, Shanghai, China.*

Zhijie Xia *Shanghai Huashan Hospital, Fudan University, China.*
janexia2006@126.com

Asian medicine: protect rare plants

As the global market in traditional Chinese medicines expands, many wild plants are on the brink of extinction (see also *Nature* **481**, 265; 2012). Urgent measures must be taken to ensure that these rare plants are harvested sustainably.

Some 11,000 plant species are listed in the Chinese pharmacopoeia, medicinal botany textbooks and ancient Chinese medical texts such as the *Compendium of Materia Medica* and *Shennong's Classic*

of *Materia Medica*.

Examples of critically depleted natural populations include *Herba epimedii*, a herb used as an aphrodisiac, tonic and antirheumatic in China, Korea and Japan; *Panax ginseng*, a tonic and sleep-inducer; *Euchresta japonica*, for anti-tumour activity; *Dysosma versipellis*, a cleanser of toxins; and *Aconitum brachypodium*, an anti-inflammatory.

Hua-Feng Zhang *Shaanxi Normal University, Xi'an, Shaanxi, China.*

Xiao-Hua Yang *Xi'an Jiaotong University, Xi'an, Shaanxi, China; and Moscow State University, Russia.*
yxh@bk.ru

Safety-test initiatives for nanomaterials

Your report on the need to establish safety regulations for nanomaterials focuses largely on US initiatives (*Nature* **480**, 160–161; 2011). Other initiatives are also making important contributions.

The Organisation for Economic Co-operation and Development (OECD) provides guidance on what parameters should be used for reporting the safety testing of nanomaterials (see go.nature.com/yiaxnd). Projects set up to aid implementation of Europe's REACH (for 'registration, evaluation, authorisation and restriction of chemicals') legislation advise on how to review information on nanomaterials.

The European Food Safety Authority published guidance last year on risk assessment of nanotechnologies in the food chain (see go.nature.com/7131fo). The European Commission's Joint Research Centre has also set up a repository of representative nanomaterials samples.

Juan Riego-Sintes *European Commission Joint Research Centre, Institute for Health and Consumer Protection, Ispra, Italy.*
juan.riego-sintes@ec.europa.eu

Philip D. Lawley

(1927–2011)

Chemist who discovered that cancer is caused by damage to DNA.

Philip Lawley laid the foundation for the idea that cancer is a genetic disease. He provided the first convincing evidence that DNA is the key target for chemicals that cause cancer, and identified a major DNA-repair mechanism that counteracts the assault of carcinogens on DNA.

Lawley, who died on 18 December 2011, was born in 1927 in the village of Abbots Bromley in Staffordshire, UK, to two teachers. He obtained a degree in chemistry in 1949 at the University of Oxford and moved to the University of Nottingham, where he studied the physicochemical properties of DNA (before its double-helical structure was revealed by James Watson and Francis Crick). His graduate work included the first investigation of what later became known as DNA denaturation (the separation of the macromolecule into two strands).

After gaining a PhD in 1953, Lawley took up a post at the Chester Beatty Research Institute (now the Institute of Cancer Research) in London, where he investigated the interaction of anticancer drugs with DNA. By 1957, he had demonstrated that alkylating agents could bond covalently with DNA to produce stable products (adducts), a radical idea at a time when it was believed that such interactions were weak or reversible and that proteins were the crucial target of carcinogens.

In the late 1950s, with chemist Peter Brookes, Lawley carried out seminal studies on the binding of mutagens and carcinogens to DNA in living organisms. The Radiochemical Centre in Amersham had just begun producing radioactively labelled chemicals, including mustard gas, which had been used as a chemical weapon in the First World War. Ironically, Lawley's father had been exposed to it while serving in the trenches. Mustard gas was already known to be a mutagen and it was later found to cause lung cancer. It was the archetype of difunctional (having two chemically reactive sites) anticancer drugs, such as melphalan and chlorambucil, which were developed during the 1950s at the Chester Beatty.

To conduct these potentially hazardous studies, Lawley and Brookes moved out of London to the institute's Pollards Wood Research Station, which was set in 24 hectares

of woodland and conveniently close to the Radiochemical Centre. In these leafy and peaceful surroundings, they produced their seminal work on the binding of mutagens and carcinogens to DNA. In 1960, they reported that mustard gas formed adducts with guanine in DNA, in viruses, bacteria, cultured mouse cancer cells and cancers in mice.

The adduct hindered DNA replication and cell division by crosslinking the strands of the DNA helix, thus explaining mustard gas's extraordinary toxicity to cells. They demonstrated that bacteria resistant to

when metabolically activated. Lawley and Brooks found that the amount of binding to DNA, but not to protein or RNA, increased with the carcinogenic potency of the PAHs. This discovery overturned the prevailing view that proteins were the critical cellular targets for carcinogens and it changed the course of cancer research.

Lawley went on to show how point mutations are induced when potent alkylating mutagens (such as *N*-methyl-*N*-nitrosourea) react with those atoms in DNA that determine base-pairing during DNA replication.

Point mutations are now known to occur frequently in a variety of human cancer genes. He also discovered the 6-*O*-methylguanine-DNA methyltransferase DNA-repair system that selectively removes such modified DNA bases.

Outwardly shy, Lawley had a wry sense of humour, liberal views and enormous integrity. He did not follow scientific fashion and was happiest working at the bench. He took extraordinary pleasure in conversation, which was enriched by his passions for books, jazz, art, dogs and old trams. He often engaged in protracted and sometimes heated debate with whoever happened to be around — usually a scientific colleague, but occasionally an innocent bystander — sometimes losing track of time. At Pollards Wood, these conversations would often continue *al fresco* in the coppiced beech woods, sometimes in the company of Lawley's whippet.

Lawley retired in 1992, but continued to work in the Institute of Cancer Research laboratories at Sutton in south London for many more years, enhancing the lives of those who worked alongside him with his knowledge, congeniality and wit. Fittingly, in January 2003, the institute honoured Brookes and Lawley by naming a £21-million (US\$32-million) laboratory in Sutton after them. ■

Stanley Venitt is honorary associate and **David H. Phillips** is professor of environmental carcinogenesis in the Analytical and Environmental Sciences Division, King's College, London SE1 9NH, UK. Both were colleagues of Philip Lawley at the Institute of Cancer Research. e-mail: stan.venitt@kcl.ac.uk



F. L. MARTIN

mustard gas used enzymes to remove those crosslinks, but sensitive bacteria did not. This was the first convincing evidence for the repair of DNA adducts in a living organism. It is now clear, 50 years on, that many of the genes mutated in human cancer are involved in DNA repair.

In a second landmark paper in 1964, Lawley and Brookes reported in *Nature* a correlation between carcinogenicity and DNA binding, using radioactively labelled polycyclic aromatic hydrocarbons (PAHs) applied to mouse skin. These ubiquitous products of incomplete combustion (found, for example, in tobacco smoke) are chemically inert, but many are potent carcinogens

element is studied. Druitt and co-workers¹ measured the chemical gradients of magnesium across individual zones in feldspar crystals. Magnesium diffuses distances of about 10 micrometres a year at temperatures of 855 °C, and so is ideal for investigating processes that happen on timescales of centuries or less. The authors used the chemical gradients of the very slow-diffusing trace elements strontium and titanium to establish that the crystal zones they studied have a thermal origin, resulting from the addition of hotter magmas to the reservoir. They propose that a significant volume of this hot magma was added in one or more spurts during the decades or months that preceded the Minoan eruption. Each spurt would have generated vigorous seismic activity and considerable inflation of the island's surface, probably serving as harbingers of catastrophe for Santorini's Minoan inhabitants³.

This short timescale between recharge of the reservoir with hot magma and subsequent eruption, compared with the 18,000-year period of dormancy between the Minoan eruption and the previous caldera-forming eruption at Santorini, echoes the short lead-in time previously calculated⁴ for the much smaller 1925–28 lava eruptions on Nea Kameni island in the centre of Santorini's caldera (Fig. 1). The net pre-Minoan influx rate of hot magma calculated by Druitt *et al.*¹ is in excess of 0.05 km³ yr⁻¹, some 50 times greater than the long-term average rate at Santorini. This is consistent with an emerging consensus that assembly of magma reservoirs is piecemeal and pulsatory, and may accelerate before eruption. These pulses can potentially be modelled from real-time ground deformation². The magnitude of deformation depends on the depth and shape of the magma reservoir, the size of the pulses, and how the heated rocks above the magma reservoir deform.

Santorini seems to be a volcano of habit. The island has seen two major cycles of magmatic activity, each lasting about 180,000 years and culminating in two large caldera-forming eruptions of silica-rich rhyodacite magma about 20,000 years apart; the Minoan is the most recent of these⁵. The early stages of volcanic activity in each cycle were marked by much smaller eruptions of magma that was more silica-poor (dacites, andesites and basalts). Since at least 197 BC, Santorini's caldera has been refilling with dacite lavas and the products of associated minor explosive eruptions⁵; the most recent lava eruption was on Nea Kameni in 1950. These lavas bear chemical similarities to the hotter magmas that flooded the magma reservoir before the Minoan eruption¹. It is not clear whether they herald the start of a new eruptive cycle on Santorini or are vestiges of the last one.

Since early 2011, there have been increasing numbers of small earthquakes beneath

the floor of the Santorini caldera⁶, several centimetres of surface uplift⁷ and increased release of magmatic carbon dioxide from the ground⁸. So far it is not certain whether these are the signs of an impending eruption — which would probably produce a dacite lava flow similar to the small, post-Minoan eruptions on Nea Kameni — or the usual restlessness that characterizes many dormant volcanoes. The difficult task facing those charged with monitoring Santorini⁶ is interpreting the ongoing signs of unrest in the context of the most likely magmatic and tectonic processes responsible and the long-term behaviour of the volcano. Applications of diffusion chronometry to volcanic rocks from Santorini^{1,4} provide valuable insight into the magnitude and timing of subterranean

magmatic events that preceded previous eruptions both large and small. ■

Jon Blundy and Alison Rust are at the School of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK.

e-mail: jon.blundy@bris.ac.uk

1. Druitt, T. H., Costa, F., Deloule, E., Dungan, M. & Scaillet, B. *Nature* **482**, 77–80 (2012).
2. Sparks, R. S. J. *Earth Planet. Sci. Lett.* **210**, 1–15 (2003).
3. Cioni, R., Gurioli, L., Sbrana, A. & Vougioukalakis, G. *Phys. Chem. Earth A* **25**, 719–724 (2000).
4. Martin, V. M. *et al. Science* **321**, 1178 (2008).
5. Druitt, T. H. *et al. Santorini Volcano Geol. Soc. Lond. Mem.* **19** (1999).
6. <http://ismosav.santorini.net>
7. Newman, A. V. *et al. Eos Trans. AGU* **92**, Fall Meet. Suppl. Abstr. V53C-2643 (2011).
8. Parks, M. *et al. Eos Trans. AGU* **92**, Fall Meet. Suppl. Abstr. V53E-2679 (2011).

NEUROSCIENCE

Reward alters specific connections

How does the brain couple a fleeting sensory input to a delayed reward during learning? A study in locusts shows that coincident firing of neurons can 'mark' a neuronal connection for later modulation. [SEE ARTICLE P.47](#)

TIMOTHY E. HOLY

A whiff of food followed by a tasty meal can trigger a specific learned association. In the quest to explain learning mechanistically, a long-standing hypothesis known as Hebb's rule^{1,2} posits that the pairing of stimulus and reward causes their respective neurons to be simultaneously active, and that this synchronous activity then changes the strength of the connections — synapses — between the neurons. But when stimulus and reward are not simultaneous, how does the brain pick the right synapses to modify? On page 47 of this issue, Cassenaer and Laurent³ report that correlated neuronal activity, when followed by delivery of a neurotransmitter molecule known to be released by reward, changes synapses in distinctive ways in the locust olfactory system. Moreover, the authors observed these changes in live insects and using natural stimuli (odours).

The relevance of Hebb's rule in associative learning received significant experimental support with the discovery of long-term potentiation⁴, in which coincident firing of two interconnected neurons results in a stronger connection between them. However, perhaps the clearest correlate of associative learning came much later, with the discovery of spike-timing-dependent plasticity (STDP)⁵. In this process, the firing (voltage spikes) of

the pre- and postsynaptic neurons must occur within a few tens of milliseconds for synaptic change (plasticity) to occur.

STDP has been extensively studied and is widespread in the brain. However, any attempt to explicitly relate this cellular phenomenon to behavioural learning faces numerous obstacles. A simplistic view specifies that the presynaptic neurons should be stimulus responsive, and the postsynaptic neurons reward responsive. However, one must acknowledge that a dizzyingly large number of arbitrary stimuli can be associated with a tiny number of intrinsic rewards. Moreover, Hebb's rule requires that all possible connections are initially present so that they might be strengthened or weakened through experience. The result might be a 'reward bottleneck', in which numerous stimulus-responsive neurons compete with each other for wiring space and/or influence over a few reward-responsive neurons.

A second obstacle arises because neurons in the brain display ongoing 'spontaneous' activity, and accidental coincidences between the spikes of pre- and postsynaptic neurons would also trigger plasticity. So, the robustness of synaptic change must be considered in light of this 'noise' problem. The narrow time window for STDP reduces the likelihood of accidental coincidence, but the sheer number of connections made by typical presynaptic

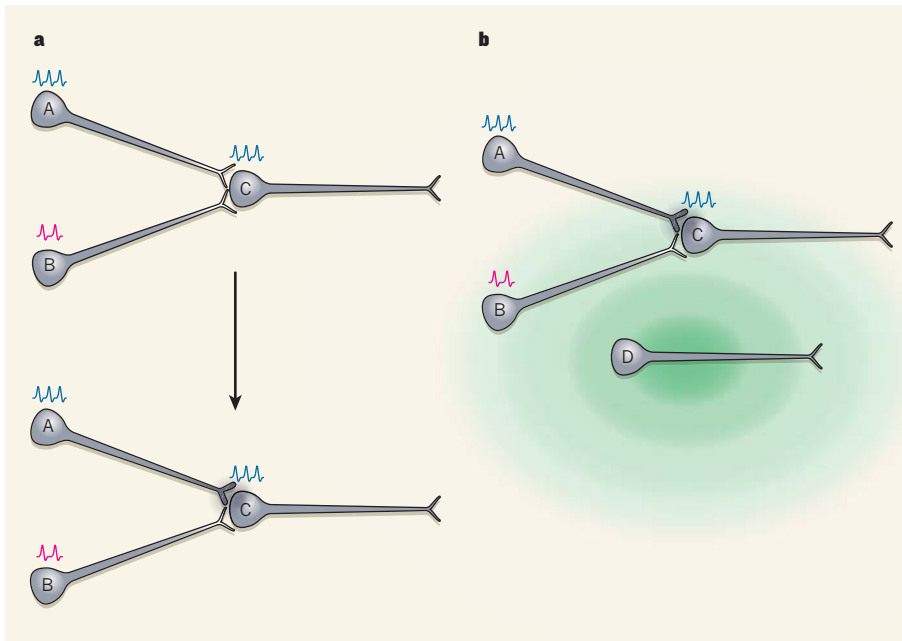


Figure 1 | Modulation of neuronal connections. Cassenaer and Laurent³ observed synaptic modulation in a specific circuit of the olfactory system of live locusts. A and B represent neurons known as Kenyon cells, and C is a neuron from another part of the brain, the β -lobe. **a**, The strength of the synaptic connection between A and C changes when both cells fire repeatedly (blue spikes) within tens of milliseconds of each other. This process, known as spike-timing-dependent plasticity (STDP), may underlie some forms of associative learning in other systems, especially if stimulus and reward are simultaneous and are sensed by neurons such as A and C. The connection between B and C, by contrast, is not altered because they do not fire coincidentally (red and blue spikes). **b**, STDP alone, however, cannot explain how stimulus and reward are connected when they occur hundreds of milliseconds apart. A possible missing link may be neuromodulatory transmitters (green) that are secreted by certain neurons (such as the reward-responsive neuron D) throughout the brain. The authors³ show that one of these molecules, octopamine, can specifically weaken the synapses (such as the A–C synapse) that had been ‘marked’ 1 second earlier by STDP.

neurons seems to suggest⁶ that random events are too common for reliable and stable learning through STDP alone.

A third problem is that the time window of STDP is too short to explain the seconds-long delays between stimulus and reward in typical associative-learning experiments. In some cases, sustained activity of presynaptic neurons⁷ may circumvent this delay issue, but there are examples for which this does not seem to be the case. A particularly relevant example is found in the insect olfactory system, in which a group of neurons called Kenyon cells, residing in a part of the brain known as the mushroom body, respond to odours too transiently for their response to coincide with the delayed delivery of reward⁸.

For the above and other reasons, researchers have focused on additional factors that regulate synaptic plasticity. Among the most prominent examples are neuromodulatory transmitters, such as acetylcholine, serotonin, dopamine, noradrenaline and octopamine. These molecules can increase or decrease the strength of synapses, and are therefore attractive candidates for mediating plasticity and learning. In both vertebrates and invertebrates, a modest number of modulatory neurons release these transmitters throughout the brain.

By making reward-coupled signals widely available without the requirement for direct synapses onto reward-activated neurons, this architecture seems to relieve the reward bottleneck. However, it also implies that neuromodulators themselves cannot encode specific experiences. An attractive solution would be an interaction between neuromodulators and a synapse-specific process such as STDP. Indeed, numerous *in vitro* studies have revealed that the two do interact⁹, but data on the relevance of these interactions *in vivo*, using natural sensory input, seem to be scarce.

Enter Cassenaer and Laurent³. Using live locusts as model organisms, they studied the synapses between the Kenyon cells and a small collection of neurons located in another part of the brain, the β -lobe. Unlike the sparsely firing Kenyon cells, β -lobe neurons fire in response to nearly every odour presented to the insect. By inserting tiny wires into the locust brain, the authors could stimulate different Kenyon cells with precise timing and record the corresponding inputs received by particular β -lobe neurons. Those synapses in which the inputs paired with the spikes of the β -lobe neuron became stronger or weaker in a typical STDP fashion (Fig. 1a).

In insects, the most well-studied mediator

of reward is octopamine, and receptors for this neurotransmitter are abundant in the β -lobe^{10,11}. To mimic associative-learning experiments (in which there is a delay between stimulus and reward), the authors³ delivered a brief injection of octopamine one second after spike pairing. They found that this neurotransmitter dramatically changed how STDP worked: the synapse invariably became weaker, even under conditions in which it would normally have grown stronger. That is, plasticity still depended on precise temporal pairing of pre- and postsynaptic activity, but the addition of octopamine altered the outcome (Fig. 1b). The researchers obtained similar results when they used odour delivery, instead of electrical stimulation, to stimulate Kenyon cells.

These changes can be considered associative only if they are specific to the particular sensory experience paired with reward. In a beautiful final experiment, Cassenaer and Laurent³ presented different odours sequentially to the locust, occasionally pairing one of them with octopamine injection. Whereas the neuronal responses to unpaired odours were scarcely affected by the neurotransmitter, the response to the paired odour was markedly decreased. Consequently, plasticity depended on the conjunction of odour-triggered activity and octopamine. It is worth noting that noradrenaline, an octopamine-related molecule, also changes olfactory circuitry in mammals and can trigger learning when paired with an odour^{12,13}. Cassenaer and Laurent reveal the interaction of synapse-specific coincident firing and nonspecific neuromodulation — a combination that may help to resolve some major challenges to Hebbian plasticity, especially the delay problem.

Given the inherent complexities in relating brain-circuit processes to behavioural learning, many questions remain. One fundamental issue is whether octopamine resolves the noise problem of accidental coincidence. Indeed, over the short timescales of Cassenaer and Laurent's experiments³, octopamine-modulated STDP was not obviously stronger, more quickly induced or more persistent than classical STDP. How, then, does (rare) neuromodulatory STDP overcome or overwhelm (common) ongoing STDP? The authors argue that, under normal conditions, ongoing STDP is balanced by mutually inhibitory activities of postsynaptic neurons, and so it requires new conditions — such as the presence of octopamine — to produce changes in neuronal firing. However, in the face of ongoing STDP, can such changes persist for long? These and similar questions will surely be tackled in future chapters of the long and complex story of memory. ■

Timothy E. Holy is in the Department of Anatomy and Neurobiology, Washington University School of Medicine in St. Louis, St. Louis, Missouri 63110, USA.
e-mail: holy@wustl.edu

1. Bain, A. *Mind and Body* (Appleton, 1877; reformatted Univ. Michigan Library, 2006).
2. Hebb, D. O. *The Organization of Behavior* (Psychology Press, 2002).
3. Cassenaer, S. & Laurent, G. *Nature* **482**, 47–52 (2012).
4. Bliss, T. V. P. & Lomo, T. J. *Physiol. (Lond.)* **232**, 331–356 (1973).
5. Markram, H., Gerstner, W. & Sjöström, P. J. *Front. Synaptic Neurosci.* **3**, 4 (2011).
6. Morrison, A., Aertsen, A. & Diesmann, M. *Neural Comput.* **19**, 1437–1467 (2007).
7. Drew, P. J. & Abbott, L. F. *Proc. Natl Acad. Sci. USA* **103**, 8876–8881 (2006).
8. Ito, I., Ong, R. C., Raman, B. & Stopfer, M. *Commun. Integr. Biol.* **1**, 170–171 (2008).
9. Pawlak, V., Wickens, J. R., Kirkwood, A. & Kerr, J. N. *Front. Synaptic Neurosci.* **2**, 146 (2010).
10. Degen, J., Gewecke, M. & Roeder, T. *Br. J. Pharmacol.* **130**, 587–594 (2000).
11. Sinakevitch, I., Mustard, J. A. & Smith, B. H. *PLoS ONE* **6**, e14536 (2011).
12. Linster, C., Nai, Q. & Ennis, M. J. *Neurophysiol.* **105**, 1432–1443 (2011).
13. Shea, S. D., Katz, L. C. & Mooney, R. J. *J. Neurosci.* **28**, 10711–10719 (2008).

SURFACE CHEMISTRY

Crystal cuts on the nanoscale

A simple method has been developed to control the shape of nanoscale cuprous oxide crystals. Some shapes turn out to be much better than others as catalysts for a light-activated reaction.

PEIDONG YANG

The most popular diamond cut is the ‘modern round brilliant’, in which a gem is hewn to expose a specific set of crystal facets so as to produce exceptional brilliance. The arrangement of the facets is crucial for providing the maximum light return through the top of the diamond. It is becoming increasingly clear that crystal shape at the nanometre scale is similarly important for getting the best performance from solid catalysts. Writing in the *Journal of the American Chemical Society*, Huang *et al.*¹ report a beautiful example of how the specific crystal facets on nanoparticles of a material affect the catalytic properties of the particles in light-activated reactions.

Cuprous oxide (Cu_2O) has commonly been used as a pigment, a fungicide and an antifouling agent for marine paints. More recently, it has been reintroduced as a promising material for the cathode of photoelectrochemical cells, which use sunlight to split water into hydrogen and oxygen². Huang *et al.*¹ have discovered a simple process for controllably synthesizing cuprous oxide nanocrystals as cubes, rhombic dodecahedra (polyhedra with 12 rhombic faces) and a variety of geometries in between (Fig. 1). The rhombic dodecahedra expose only one kind of facet (known in crystallographic jargon as {110} facets), and the authors observed that these nanocrystals are exceptionally good catalysts for the light-induced degradation of the organic compound methyl orange. This reaction serves as a model for reactions involved in clearing organic pollutants from the environment.

One of the tricks often used to control which shapes and facets form during the synthesis

of nanocrystals is to preferentially adsorb certain molecular species to specific planes of atoms exposed on crystal surfaces. This strategy generates direction-dependent nanocrystal growth by stabilizing a particular facet — growth is limited on crystal planes where

adsorption (and so stabilization) is strong, but is promoted on planes where adsorption is weak. A wide variety of chemical species can facilitate shape control in the growth of metal and oxide nanocrystals, including: surfactants, polymers and biomolecules; small molecules such as those in gases; and even atomic species, such as anions or metal ions.

It has previously been reported³ that the shape of cuprous oxide nanocrystals can be controlled by varying the ratio of chloride ions to sodium dodecyl sulphate ions in the solutions from which the crystals form; the chloride ions stabilize the {100} planes of the crystals, whereas the dodecyl sulphate ions stabilize the {111} planes. In this way, a range of nanoscale polyhedra has been produced.

To make their cuprous oxide nanocrystals, Huang *et al.*¹ used a different recipe from that previously reported³, and found that, by simply adjusting the amount of the reducing agent in the reaction mixture, they could control the shapes of the crystals. Their approach allowed them to make an alternative series of polyhedra (Fig. 1) to that produced in the earlier work. In addition, their shape-control method can be used at room temperature, unlike many previously reported strategies, which typically require higher temperatures.

Different nanocrystal shapes naturally display crystallographically distinct facets. In the absence of any major surface reconstruction processes, each facet type contains surface

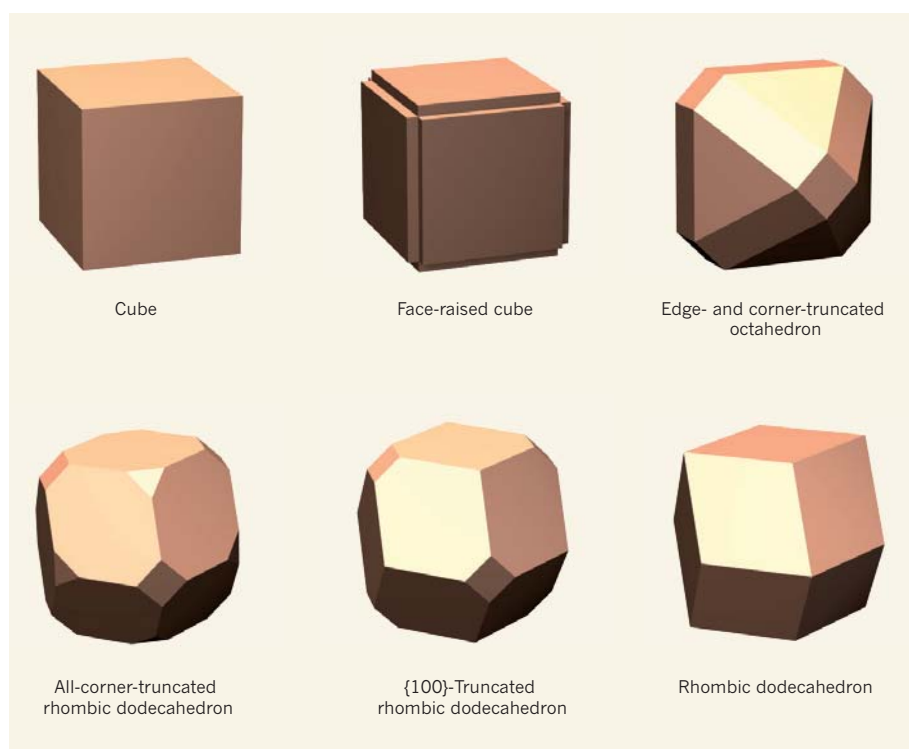


Figure 1 | Nanocrystal shapes. Huang *et al.*¹ report a simple process for controlling the shape of cuprous oxide nanocrystals during synthesis; the shapes made are depicted. The authors observed that the catalytic properties of the nanocrystals in a light-activated reaction — the degradation of methyl orange — largely depend on the kind of facets that the crystals have. Cubic crystals are essentially inactive, octahedra are moderately active, and rhombic dodecahedra are exceptionally active.

1. Bain, A. *Mind and Body* (Appleton, 1877; reformatted Univ. Michigan Library, 2006).
2. Hebb, D. O. *The Organization of Behavior* (Psychology Press, 2002).
3. Cassenaer, S. & Laurent, G. *Nature* **482**, 47–52 (2012).
4. Bliss, T. V. P. & Lomo, T. J. *Physiol. (Lond.)* **232**, 331–356 (1973).
5. Markram, H., Gerstner, W. & Sjöström, P. J. *Front. Synaptic Neurosci.* **3**, 4 (2011).
6. Morrison, A., Aertsen, A. & Diesmann, M. *Neural Comput.* **19**, 1437–1467 (2007).
7. Drew, P. J. & Abbott, L. F. *Proc. Natl Acad. Sci. USA* **103**, 8876–8881 (2006).
8. Ito, I., Ong, R. C., Raman, B. & Stopfer, M. *Commun. Integr. Biol.* **1**, 170–171 (2008).
9. Pawlak, V., Wickens, J. R., Kirkwood, A. & Kerr, J. N. *Front. Synaptic Neurosci.* **2**, 146 (2010).
10. Degen, J., Gewecke, M. & Roeder, T. *Br. J. Pharmacol.* **130**, 587–594 (2000).
11. Sinakevitch, I., Mustard, J. A. & Smith, B. H. *PLoS ONE* **6**, e14536 (2011).
12. Linster, C., Nai, Q. & Ennis, M. J. *Neurophysiol.* **105**, 1432–1443 (2011).
13. Shea, S. D., Katz, L. C. & Mooney, R. J. *J. Neurosci.* **28**, 10711–10719 (2008).

SURFACE CHEMISTRY

Crystal cuts on the nanoscale

A simple method has been developed to control the shape of nanoscale cuprous oxide crystals. Some shapes turn out to be much better than others as catalysts for a light-activated reaction.

PEIDONG YANG

The most popular diamond cut is the ‘modern round brilliant’, in which a gem is hewn to expose a specific set of crystal facets so as to produce exceptional brilliance. The arrangement of the facets is crucial for providing the maximum light return through the top of the diamond. It is becoming increasingly clear that crystal shape at the nanometre scale is similarly important for getting the best performance from solid catalysts. Writing in the *Journal of the American Chemical Society*, Huang *et al.*¹ report a beautiful example of how the specific crystal facets on nanoparticles of a material affect the catalytic properties of the particles in light-activated reactions.

Cuprous oxide (Cu_2O) has commonly been used as a pigment, a fungicide and an antifouling agent for marine paints. More recently, it has been reintroduced as a promising material for the cathode of photoelectrochemical cells, which use sunlight to split water into hydrogen and oxygen². Huang *et al.*¹ have discovered a simple process for controllably synthesizing cuprous oxide nanocrystals as cubes, rhombic dodecahedra (polyhedra with 12 rhombic faces) and a variety of geometries in between (Fig. 1). The rhombic dodecahedra expose only one kind of facet (known in crystallographic jargon as {110} facets), and the authors observed that these nanocrystals are exceptionally good catalysts for the light-induced degradation of the organic compound methyl orange. This reaction serves as a model for reactions involved in clearing organic pollutants from the environment.

One of the tricks often used to control which shapes and facets form during the synthesis

of nanocrystals is to preferentially adsorb certain molecular species to specific planes of atoms exposed on crystal surfaces. This strategy generates direction-dependent nanocrystal growth by stabilizing a particular facet — growth is limited on crystal planes where

adsorption (and so stabilization) is strong, but is promoted on planes where adsorption is weak. A wide variety of chemical species can facilitate shape control in the growth of metal and oxide nanocrystals, including: surfactants, polymers and biomolecules; small molecules such as those in gases; and even atomic species, such as anions or metal ions.

It has previously been reported³ that the shape of cuprous oxide nanocrystals can be controlled by varying the ratio of chloride ions to sodium dodecyl sulphate ions in the solutions from which the crystals form; the chloride ions stabilize the {100} planes of the crystals, whereas the dodecyl sulphate ions stabilize the {111} planes. In this way, a range of nanoscale polyhedra has been produced.

To make their cuprous oxide nanocrystals, Huang *et al.*¹ used a different recipe from that previously reported³, and found that, by simply adjusting the amount of the reducing agent in the reaction mixture, they could control the shapes of the crystals. Their approach allowed them to make an alternative series of polyhedra (Fig. 1) to that produced in the earlier work. In addition, their shape-control method can be used at room temperature, unlike many previously reported strategies, which typically require higher temperatures.

Different nanocrystal shapes naturally display crystallographically distinct facets. In the absence of any major surface reconstruction processes, each facet type contains surface

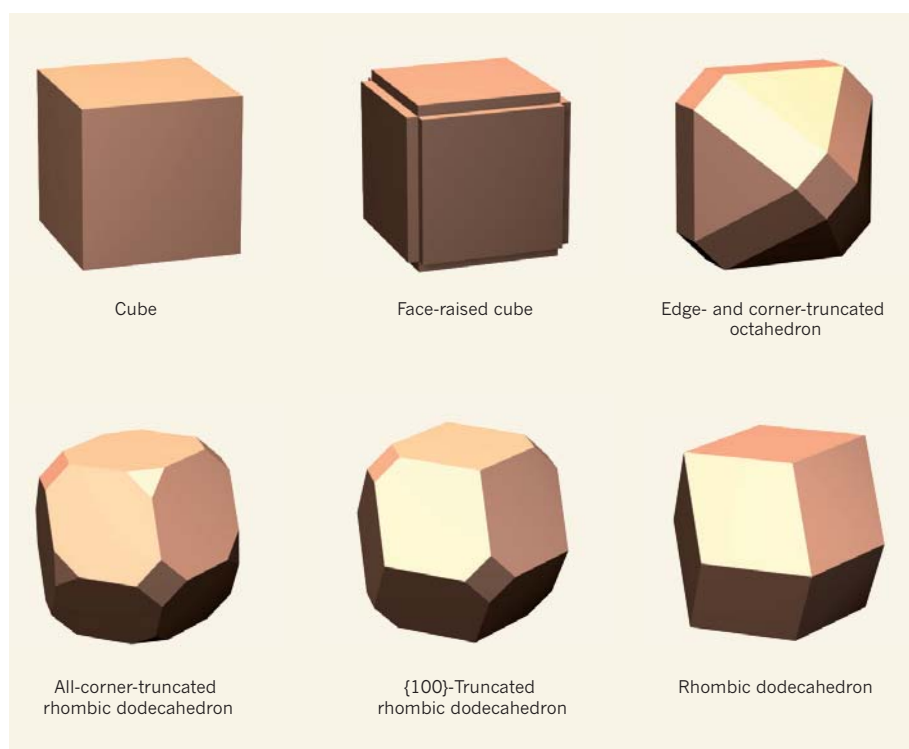


Figure 1 | Nanocrystal shapes. Huang *et al.*¹ report a simple process for controlling the shape of cuprous oxide nanocrystals during synthesis; the shapes made are depicted. The authors observed that the catalytic properties of the nanocrystals in a light-activated reaction — the degradation of methyl orange — largely depend on the kind of facets that the crystals have. Cubic crystals are essentially inactive, octahedra are moderately active, and rhombic dodecahedra are exceptionally active.

atoms that reside in particular environments (arrangements of nearby atoms). This in turn causes the surface atoms in different facets to have distinct catalytic activities. This fundamental insight dates back to the classic discovery that {111} surfaces on single crystals of platinum are typically much more catalytically active than {100} surfaces in reactions that form aromatic compounds.

Indeed, Huang and colleagues' cuprous oxide nanocrystals clearly exhibit very different catalytic properties depending on the facets displayed. The authors found that cubic nanocrystals that are bound exclusively by {100} faces are essentially inactive in the photodegradation reaction of methyl orange, whereas octahedra bound by {111} surfaces are moderately active catalysts. Rhombic dodecahedra bound by {110} surfaces, however, caused much faster photodegradation of methyl orange than the other nanocrystal shapes. The authors speculate that the observed catalytic activities correlate with the surface density of copper sites on the different facets, although they report no direct evidence for this. It is also unclear whether many of the atoms on nanocrystal surfaces undergo rearrangements during the reactions, or how ligand molecules that are bound to the different surfaces — either during crystal formation or in the photodegradation reactions — affect catalytic activity.

The new findings¹ form part of a growing body of literature documenting surface-dependent catalytic activities, perhaps most notably for metals. For example, key studies performed on the surfaces of samples of macroscopic single crystals under ultra-high-vacuum conditions have shown that some reactions are sensitive to surface structure⁴. A quintessential case is the platinum-catalysed reaction of benzene with hydrogen: the {100} platinum surface yields only cyclohexane (a saturated hydrocarbon) as a product, whereas the {111} surface also yields cyclohexene (an unsaturated analogue of cyclohexane)⁴.

This dependence of catalytic activity on surface structure extends to metal nanocrystals, particularly in the case of platinum, palladium and rhodium. For instance, in the hydrogenation of benzene, platinum nanocrystals that expose well-defined {100} and {111} surfaces behave⁵ in much the same way as larger crystals. In another example, faceted palladium nanocrystals that have high-index surfaces — {730} and {221} surfaces, which contain a high density of atomic steps and ledges — were found⁶ to be much more active catalysts in 'Suzuki' carbon-carbon bond-formation reactions than nanoparticles that have typical {100} surfaces. Such high-index faceted nanoparticles are also better catalysts for some electrochemical reactions^{7,8}. And returning to cuprous oxide, highly faceted polyhedral microcrystals that display high-index {311} surfaces are especially good catalysts for the oxidation of carbon monoxide⁹.

Although it is clear that the reactivity and selectivity of nanoparticle catalysts depend on the shape (and hence the exposed surfaces) of the particles, in many cases it is unclear whether these effects are truly surface-dependent. This is especially true for reactions in solution or in the gas phase, where there is evidence that changes to particle surfaces and shapes may occur during the reactions¹⁰. Methods for the direct *in situ* surface analysis of molecules adsorbed to nanocrystals in solution, analogous to the techniques commonly used to study larger single crystals, need to be developed to address this issue. Nevertheless, as Huang *et al.*¹ have shown, sculpting the shape of nanocrystals is a promising approach for developing catalysts that produce only one desired reaction product out of many other possible products at high reaction rates — an essential requirement for a wide variety of industrially important reactions. ■

PLANT BIOLOGY

Equal-parenting policy

During early embryo development in animals, maternal genes are expressed in preference to those of the zygote — the newly fertilized egg. But in plants it seems that zygote genomes switch on within hours of fertilization. [SEE LETTER P.94](#)

CHRISTOPHER J. HALE
& STEVEN E. JACOBSEN

Fertilization occurs when two gametes merge to form a zygote. The zygote's genome comprises two sets of chromosomes — one maternal, the other paternal. In animals, the stages of development immediately after fertilization are disproportionately controlled by maternally inherited factors, with the zygotic genome being switched on gradually in waves of activation¹. Whether plants undergo a similar maternal-to-zygotic transition in genomic activity has been a long-standing question. In contrast to previous work², Nodine and Bartel³ show (page 94 of this issue) that a plant zygote's protein-coding transcriptome — its total complement of protein-coding RNA transcripts — contains relatively equal maternal and paternal contributions. It seems, therefore, that plants and animals have evolved distinct strategies for managing the early steps in the transition from two gametes to an organism with maternal and paternal copies of the genome⁴.

For the plant species *Arabidopsis thaliana*, a wealth of whole-genome sequence data exists for closely related yet genetically distinct lines known as ecotypes⁴. To directly measure the respective contributions of the maternal

Peidong Yang is in the Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, USA, and at King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: p_yang@berkeley.edu

- Huang, W.-C., Lyu, L.-M., Yang, Y.-C. & Huang, M. H. *J. Am. Chem. Soc.* **134**, 1261–1267 (2012).
- Paracchino, A., Laporte, V., Sivula, K., Grätzel, M. & Thimsen, E. *Nature Mater.* **10**, 456–461 (2011).
- Siegfried, M. J. & Choi, K.-S. *Adv. Mater.* **16**, 1743–1746 (2004).
- Somorjai, G. A. & Li, Y. *Introduction to Surface Chemistry and Catalysis* 2nd edn, Ch. 9 (Wiley, 1994).
- Bratlief, K. M., Lee, H., Komvopoulos, K., Yang, P. & Somorjai, G. A. *Nano Lett.* **7**, 3097–3101 (2007).
- Wang, F. *et al.* *J. Am. Chem. Soc.* **133**, 1106–1111 (2011).
- Chen, Q.-S. *et al.* *J. Am. Chem. Soc.* **133**, 12930–12933 (2011).
- Habas, S. E., Lee, H., Radmilovic, V., Somorjai, G. A. & Yan, P. *Nature Mater.* **6**, 692–697 (2007).
- Leng, M. *et al.* *J. Am. Chem. Soc.* **132**, 17084–17087 (2010).
- Joo, S. *et al.* *Nature Mater.* **8**, 126–131 (2009).

and paternal genomes to the zygotic transcriptome, Nodine and Bartel crossed two of these ecotypes and sequenced the RNA content of the resulting zygote. Using the genetic differences between the parents, they then matched these RNAs to either the maternal or the paternal genome.

The authors observed a near-equal abundance of RNA sequences derived from each parental genome in the plant embryos as early as the one- to two-cell stage. Only a small number of genes showed biased expression. They also performed reciprocal crosses — switching which ecotype provided the male and female gamete. It emerged that most of these cases of bias were probably due to genetic or epigenetic differences (chemical modifications that alter gene expression without affecting the DNA sequence) between the two ecotypes used, rather than being a result of which ecotype was used as the male or the female parent. These findings imply that the plant zygotic genome is essentially switched on only hours after fertilization.

Nodine and Bartel's results are in contrast to other reports^{2,5} suggesting that the maternal-to-zygotic transition in plants is gradual, as is the case for animals. The most comprehensive of these reports², published last year, used a genomics approach similar to that of the current study, and found that more than 80% of the *Arabidopsis* transcriptome in early

⁴This article and the paper³ under discussion were published online on 22 January 2012.

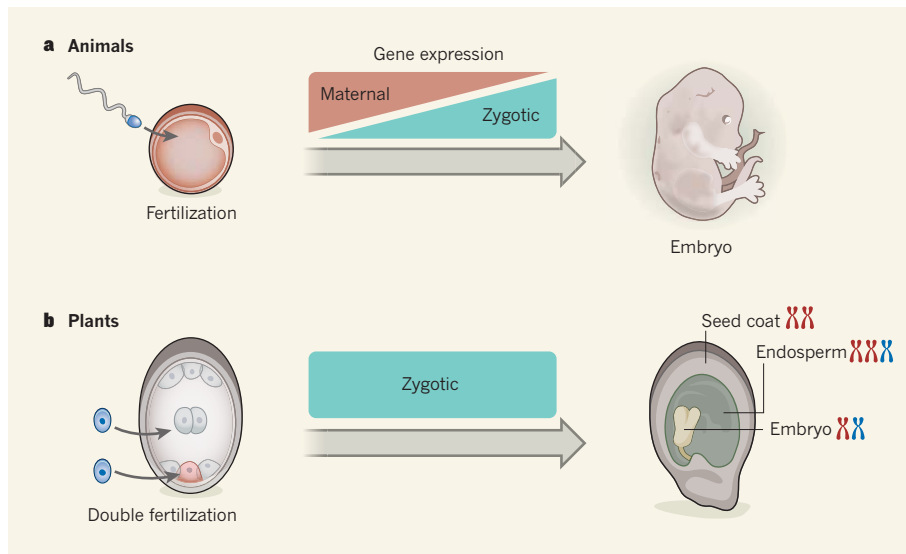


Figure 1 | Maternal-to-zygotic genomic transition in plants and animals. **a**, In animals, a single fertilization event between maternal and paternal gametes forms the zygote (not shown). Initially, the animal embryo is under predominant control of the maternal genome, and the zygotic genome is only gradually activated over the course of embryogenesis. **b**, In plants, a double fertilization process generates the zygote and the endosperm tissue. Both the endosperm and the embryo are encased in maternal tissue that generates the seed coat. Chromosome symbols represent maternal (red) and paternal (blue) genomes. Nodine and Bartel³ show that, in contrast to the case for animals, the plant zygotic genome is activated almost immediately after fertilization.

embryos is derived from the maternal genome.

What might explain such different results between the two studies? Nodine and Bartel³ suggest that the most likely answer is that the embryo samples of the earlier study were contaminated by maternal tissue. This is because, in *Arabidopsis* as in other flowering plants, the embryo is surrounded by a seed coat and other maternally derived tissue (Fig. 1), making the isolation of high-purity embryonic tissue difficult — a problem for the extremely sensitive genome-sequencing techniques used in both studies. After observing this maternal-contamination effect in pilot studies, Nodine and Bartel³ overcame the problem by extensively washing the isolated embryo cells.

One question arising from Nodine and Bartel's study is why activation of the zygotic genome is so different between plants and animals. Because the two kinds of organisms evolved multicellularity independently⁶, and have very different life histories, it is perhaps not surprising that the maternal-to-zygotic transition also differs between them. In plants, gametes are generated from cells derived from a pool of undifferentiated cells that are also responsible for generating structures such as leaves^{7,8}, rather than from a distinct germ-cell lineage as occurs in animals. In addition, fertilization itself is radically different in flowering plants compared with animals, with each seed being the product of two fertilization events (Fig. 1b). In this process, one fertilization event forms the endosperm, a tissue that is functionally similar to the mammalian placenta, and which contains two sets of maternal

chromosomes and one set of paternal chromosomes. A second fertilization event results in the formation of the embryo. It is noteworthy that, although Nodine and Bartel show that the embryo experiences equal transcriptome contributions from both parental genomes, gene expression in the embryo-nourishing endosperm shows extensive parental influence⁹; these effects are due to uneven nuclear 'dosage' as well as to epigenetic imprinting effects.

Another question is how the early plant embryo coordinates rapid integration of two genomes and the concomitant activation of a resulting zygotic genome. In animals, a suite of mechanisms acts to clear maternal factors, such as proteins and RNAs, from the developing zygote and to activate the zygotic genome^{1,10}. It is unclear whether similar mechanisms operate in plants. Tantalizingly, epigenetic processes such as DNA methylation and demethylation, as well as regulatory small RNA molecules, have recently been implicated⁹ in the control of both gamete and endosperm development. Future studies of these pathways may reveal mechanisms for the regulation of gene expression in embryonic plants.

Understanding how a functional plant genome so quickly emerges from two progenitor genomes is vital for understanding plant development, and for informing approaches to plant breeding and plant biotechnology. In many crop species, combining two different parental genomes can generate regular and predictable hybrid vigour, known as heterosis. In some species this vigour is obvious very early in development¹¹. Consistent with these observations, Nodine and Bartel's work³ suggests that



50 Years Ago

In a Cantor Lecture ... Dr. Tom A. Margerison described how science could be presented on television not only to the specialist audience but especially to the layman. In spite of the absolute necessity of science and technology, the 10–15 per cent of the population who guide the destiny of Britain, the professional men, the majority of teachers, the industrialists, the politicians, are almost completely ignorant about science. The bridging of this gap in this most influential part of the population is of great urgency. There are many ways in which television can help to close the gap.
From *Nature* 3 February 1962

100 Years Ago

Mr. Harding's letter ... reminds me of an experience which ... may be of sufficient interest to place upon record in these columns ... It must, I think, have been in 1866 or 1867 ... that I had occasion to go from the West to the East End of London. Starting upon my journey about 10 p.m., it began to rain soon after I left the house in Bayswater, and I opened an umbrella, which, to my surprise, became stiffer and heavier every moment, and was found on examination to be so thickly glazed over with ice that it was impossible to close it. At the same time the pavements and roadway were also becoming uniformly glazed; pedestrian movement was most difficult, and all horse traffic was suspended. Although an experience of some forty-five years ago, the impression left upon my memory is still vivid — the ludicrous sight of people carrying ponderous and rigidly frozen umbrellas which they could not close, the stream of skaters down Oxford Street and Holborn, and the silence due to the absence of vehicles, all came to mind on reading Mr. Harding's letter.
From *Nature* 1 February 1912

zygotic-genome dynamics in plants, including, perhaps, some aspects of heterosis, are established almost immediately after fertilization. ■

Christopher J. Hale is in the Department of Molecular, Cell and Developmental Biology, and **Steven E. Jacobsen** is at the Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mails: cjhale@ucla.edu; jacobsen@ucla.edu

1. Tadros, W. & Lipshitz, H. D. *Development* **136**, 3033–3042 (2009).
2. Autran, D. *et al.* *Cell* **145**, 707–719 (2011).

3. Nodine, M. D. & Bartel, D. P. *Nature* **482**, 94–97 (2012).
4. Cao, J. *et al.* *Nature Genet.* **43**, 956–963 (2011).
5. Vielle-Calzada, J.-P., Baskar, R. & Grossniklaus, U. *Nature* **404**, 91–94 (2000).
6. Grosberg, R. K. & Strathmann, R. R. *Annu. Rev. Ecol. Evol. Syst.* **38**, 621–654 (2007).
7. Sundaresan, V. & Alandete-Saez, M. *Development* **137**, 179–189 (2010).
8. Twell, D. *Sex. Plant Reprod.* **24**, 149–160 (2011).
9. Feng, S., Jacobsen, S. E. & Reik, W. *Science* **330**, 622–627 (2010).
10. Walser, C. B. & Lipshitz, H. D. *Curr. Opin. Genet. Dev.* **21**, 431–443 (2011).
11. Meyer, S., Pospisil, H. & Scholten, S. *Plant Mol. Biol.* **63**, 381–391 (2007).

MOLECULAR MOTORS

A staggering giant

The protein dynein ‘walks’ along filaments to transport various cargoes within the cell. Two studies reveal that, unlike other motor proteins, dynein’s steps are not strictly coordinated.

WILHELM J. WALTER & STEFAN DIEZ

When you are walking down the street your feet probably take alternating steps, with one foot passing the other each time. Alternatively, you might walk without letting your feet pass one other, in a limping motion. However, regardless of the stepping pattern, your walking will be highly coordinated. At the molecular scale, the motor protein dynein also ‘walks’ along filaments to carry vesicles and organelles to specific locations within a cell. But is coordination between dynein’s two ‘feet’ required for walking? Two papers — one by DeWitt *et al.*¹ in *Science* and another by Qiu *et al.*² in *Nature Structural & Molecular Biology* — independently address this question and come to the surprising conclusion that dynein uses both random and coordinated walking.

Cytoplasmic dynein is a giant, multi-subunit protein that uses the chemical energy stored in the molecule ATP to transport cargoes³. It moves along microtubules — long polymeric tubes composed of dimers of the protein tubulin, typically arranged to form 13 parallel tracks. Dynein has two ‘head’ domains linked by their respective ‘tails’, whereby the heads act as ‘feet’ for walking along the microtubules. Each head contains four binding pockets for ATP and a microtubule-binding site. Of the four ATP-binding pockets, one catalyses ATP hydrolysis to generate energy for walking; the other three are thought to be important for the regulation of dynein activity⁴.

Previous work^{5,6} using single dynein motors bound to beads revealed that during processive motility on microtubules —

the process by which a motor takes multiple steps without dissociating from the microtubule — dynein moves by 8-nanometre steps. This distance corresponds to the periodicity of the dimeric tubulin subunits along the microtubule tracks. However, little is known about the coordination of the two dynein heads during motion.

To make the dynein heads visible and study the stepping mechanism during processive motion, DeWitt *et al.*¹ and Qiu *et al.*² attached fluorescent markers to the dynein heads and used fluorescence microscopy. Moreover, to distinguish between the movement of the two

individual heads, both groups used markers that emitted fluorescence of a different colour for each head, a method previously applied^{7,8} to another motor protein, myosin V.

To facilitate the two-colour labelling, both groups generated dynein molecules in which the two subunits had been artificially linked to each other. DeWitt *et al.* attached different fluorescent quantum dots (semiconductor nanocrystals) to modified dynein heads that dimerized in the presence of the small molecule rapamycin. By contrast, Qiu *et al.* attached complementary DNA strands and different fluorescent organic molecules (less stable but smaller than quantum dots) to dynein heads, so that the DNA molecules self-associated to link both heads together. To observe the stepping mechanism, both groups^{1,2} slowed dynein’s movement by using an ATP concentration that was much lower than that normally found in the cell’s cytoplasm. From computerized image analysis, they were able to track the positions of individual dynein heads in two dimensions with an accuracy of about 3 nm.

The primary finding of both studies is that dynein has a highly variable stepping pattern (Fig. 1). For most steps, the heads moved alternately, but they passed each other only infrequently. After each step, the head-to-head distance varied widely, from less than 5 nm to 50 nm, which is indicative of uncoordinated stepping. By contrast, other motor proteins such as kinesin-1 show strictly coordinated stepping. In kinesin-1, only the lagging head can bind ATP, because of a mechanism that relies on the strain between the two heads. This mechanical coupling forces the heads to step in a strictly alternating pattern, because the hydrolysis of the ATP molecule causes the lagging head to make the next step⁹.

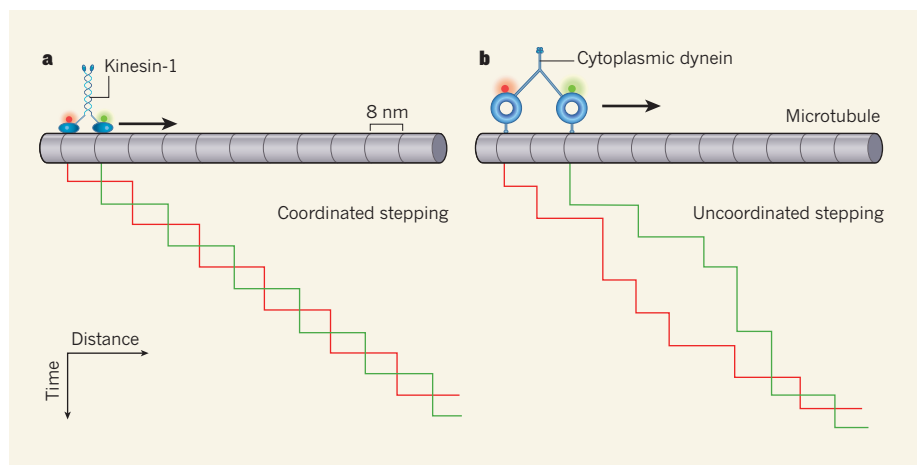


Figure 1 | Stepping patterns in motor proteins. **a**, Some motor proteins, such as kinesin-1, move in a highly coordinated way along microtubules (polymeric protein tubes; grey). The traces show the stepping of individual motor heads (each labelled with a different fluorescent marker, red or green). **b**, DeWitt *et al.*¹ and Qiu *et al.*² show that, by contrast, cytoplasmic dynein moves in a variable stepping pattern. Although the heads often move alternately, they only seldom pass each other, and the spacing between them varies strongly over time.

Interestingly, the current studies^{1,2} show that the degree of coordination in dynein stepping depends on the head-to-head distance: the stepping pattern is random when the distance is small, but becomes coordinated when the distance is large. This implies a tension-based coordination mechanism similar to the mechanical coupling seen in kinesin-1, but less pronounced.

To further test for the necessity of coordination to dynein motility, DeWitt *et al.*¹ generated a dynein motor in which one of the two heads could not hydrolyse ATP. This head associated only weakly with the microtubule track, probably because transition to a strong binding state requires ATP hydrolysis. This motor moved reliably along microtubules, however, suggesting that coordination is not required for long-range transport.

The high accuracy with which both groups^{1,2} managed to track dynein's motion on the surface of microtubules allowed them to make yet another discovery: the two heads exhibit an inherent left-right asymmetry, with the right head more likely to be in the lead. Thus, dynein can be described as staggering sideways along the microtubule.

In light of the two papers' unexpected findings, several exciting questions and ideas for follow-up experiments arise. First, the observed effect of the suggested tension-based mechanism for head coordination is rather

small; however, because both groups used modified dyneins, the effect might be more pronounced in normal dynein. It is known¹⁰ that slight modifications to the tail domains, which help to link dynein's two subunits, affect the protein's stepping pattern. Full coordination might thus rely on full-length tail domains and additional subunits, both of which are present in wild-type dynein but lacking in the artificially linked constructs analysed here.

Second, the regulatory ATP-binding sites should not be neglected. It is possible that the disordered stepping patterns observed by the authors are the result of limited ATP binding to the regulatory ATP-binding sites, because all experiments were performed at low concentrations of ATP.

And, third, the highly variable step sizes of the dynein heads — smaller than 16 nm on average — suggest that the centre of mass of the motor dimers tends to move less than 8 nm per step. But it is not clear how this can be reconciled with the aforementioned results of previous studies^{5,6} of similar dynein constructs, which were found to take 8-nm steps.

Further development of optical technologies — towards even faster image-acquisition rates at tracking accuracies similar to the outstanding ones presented by DeWitt *et al.*¹ and Qiu *et al.*² — will, step by step, allow for fascinating insight into the functioning of dynein and other molecular motors under

physiological conditions. Knowledge about the flexibility in the stepping behaviours of various motors might allow us to explain their astonishing ability to move along crowded microtubules in a dense cytoplasmic environment. ■

Wilhelm J. Walter and Stefan Diez are at the B CUBE — Center for Molecular Bioengineering, Technische Universität Dresden and at the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.

e-mails: walter@bcube-dresden.de; diez@bcube-dresden.de

1. DeWitt, M. A., Chang, A. Y., Combs, P. A. & Yildiz, A. *Science* **335**, 221–225 (2012).
2. Qiu, W. *et al.* *Nature Struct. Mol. Biol.* <http://dx.doi.org/10.1038/nsmb.2205> (2012).
3. Vallee, R. B., Williams, J. C., Varma, D. & Barnhart, L. E. *J. Neurobiol.* **58**, 189–200 (2004).
4. Sakakibara, H. & Oiwa, K. *FEBS J.* **278**, 2964–2979 (2011).
5. Toba, S., Watanabe T. M., Yamaguchi-Okimoto, L., Toyoshima, Y. Y. & Higuchi, H. *Proc. Natl Acad. Sci. USA* **103**, 5741–5745 (2006).
6. Gennerich, A., Carter, A. P., Reck-Peterson, S. L. & Vale, R. D. *Cell* **131**, 952–965 (2007).
7. Warshaw, D. M. *et al.* *Biophys. J.* **88**, L30–L32 (2005).
8. Churchman, L. S., Ökten, Z., Rock, R. S., Dawson, J. F. & Spudis, J. A. *Proc. Natl Acad. Sci. USA* **102**, 1419–1423 (2005).
9. Rosenfeld, S. S., Fordyce, P. M., Jefferson, G. M., King, P. H. & Block, S. M. *J. Biol. Chem.* **278**, 18550–18556 (2003).
10. Ori-McKenney, K. M., Xu, J., Gross, S. P. & Vallee, R. B. *Nature Cell Biol.* **12**, 1228–1234 (2010).

PRECISION MEASUREMENT

A comb in the extreme ultraviolet

A 'comb' of photons at evenly spaced frequencies in the extreme ultraviolet has been generated. It will allow a more precise search for variation in the fine-structure constant, which sets the strength of the electromagnetic force. SEE LETTER P.68

LINDA YOUNG

That improved measurement precision yields new physics is axiomatic. In the case of optical spectroscopy, the quest for ultimate precision was revolutionized around the turn of the century by the realization of the optical frequency comb^{1–3}. This elegant invention, which was honoured by the 2005 Nobel Prize in Physics^{4,5}, provides a direct link between radio and optical frequencies, allowing one to count cycles of an electromagnetic field at near-petahertz frequencies (1 petahertz is 10¹⁵ Hz). On page 68 of this issue, Cingöz *et al.*⁶ extend the frequency range of combs to the extreme-ultraviolet spectral regime.

An optical frequency comb is produced by

a mode-locked ultra-fast laser in which a light pulse of sub-picosecond duration (1 picosecond is 10^{–12} s) circulates inside an optical cavity made up of a set of highly reflective mirrors. The laser emits an infinite train of pulses at a certain repetition rate (f_{rep}) that is determined by the time the pulses take to make a round trip of the cavity's mirrors^{7,8}. This train of pulses has two distinguishing properties: the pulses are equally spaced in time, and the phase (where a wave's peaks and troughs lie) of the electromagnetic field in each pulse is precisely shifted from that in the subsequent pulse (the locked-phase relationship). The train produces a comb of sharp spectral lines uniformly spaced in frequency.

The frequencies of the comb teeth are defined by a simple formula: $f_n = nf_{\text{rep}} + f_0$,

where f_{rep} and f_0 are radio frequencies, f_n is optical frequency and n is an integer of the order of 10⁶–10⁷ (Fig. 1). With one comb tooth, f_n , referenced to an absolute standard such as the frequency of a caesium atomic clock, the absolute frequency of an unknown optical laser may be counted as a beat note relative to another comb tooth with an accuracy at the level of one part in 10¹⁴. This greatly exceeds the accuracy achievable using conventional wavelength-based measurements, which are limited by wavefront distortion to an accuracy at the 10^{–10} level.

A prominent early example of the application of the frequency comb was the measurement¹ of a two-photon transition in hydrogen to an accuracy of more than 2×10^{-14} . This measurement, combined with one performed about four years later⁹, placed a limit on the variability of the fine-structure constant — a fundamental constant that characterizes the strength of the electromagnetic interaction and whose variation in time informs the standard model of particle physics. Beyond precision spectroscopy, the optical frequency comb has fostered a multitude of applications, including a clock based on optical frequencies, massively parallel multi-wavelength spectroscopy of unknown chemicals, and long-distance transmission of ultra-stable clock signals, which are useful for next-generation telescope arrays or accelerator systems.

Interestingly, the current studies^{1,2} show that the degree of coordination in dynein stepping depends on the head-to-head distance: the stepping pattern is random when the distance is small, but becomes coordinated when the distance is large. This implies a tension-based coordination mechanism similar to the mechanical coupling seen in kinesin-1, but less pronounced.

To further test for the necessity of coordination to dynein motility, DeWitt *et al.*¹ generated a dynein motor in which one of the two heads could not hydrolyse ATP. This head associated only weakly with the microtubule track, probably because transition to a strong binding state requires ATP hydrolysis. This motor moved reliably along microtubules, however, suggesting that coordination is not required for long-range transport.

The high accuracy with which both groups^{1,2} managed to track dynein's motion on the surface of microtubules allowed them to make yet another discovery: the two heads exhibit an inherent left-right asymmetry, with the right head more likely to be in the lead. Thus, dynein can be described as staggering sideways along the microtubule.

In light of the two papers' unexpected findings, several exciting questions and ideas for follow-up experiments arise. First, the observed effect of the suggested tension-based mechanism for head coordination is rather

small; however, because both groups used modified dyneins, the effect might be more pronounced in normal dynein. It is known¹⁰ that slight modifications to the tail domains, which help to link dynein's two subunits, affect the protein's stepping pattern. Full coordination might thus rely on full-length tail domains and additional subunits, both of which are present in wild-type dynein but lacking in the artificially linked constructs analysed here.

Second, the regulatory ATP-binding sites should not be neglected. It is possible that the disordered stepping patterns observed by the authors are the result of limited ATP binding to the regulatory ATP-binding sites, because all experiments were performed at low concentrations of ATP.

And, third, the highly variable step sizes of the dynein heads — smaller than 16 nm on average — suggest that the centre of mass of the motor dimers tends to move less than 8 nm per step. But it is not clear how this can be reconciled with the aforementioned results of previous studies^{5,6} of similar dynein constructs, which were found to take 8-nm steps.

Further development of optical technologies — towards even faster image-acquisition rates at tracking accuracies similar to the outstanding ones presented by DeWitt *et al.*¹ and Qiu *et al.*² — will, step by step, allow for fascinating insight into the functioning of dynein and other molecular motors under

physiological conditions. Knowledge about the flexibility in the stepping behaviours of various motors might allow us to explain their astonishing ability to move along crowded microtubules in a dense cytoplasmic environment. ■

Wilhelm J. Walter and Stefan Diez are at the B CUBE — Center for Molecular Bioengineering, Technische Universität Dresden and at the Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.

e-mails: walter@bcube-dresden.de; diez@bcube-dresden.de

1. DeWitt, M. A., Chang, A. Y., Combs, P. A. & Yildiz, A. *Science* **335**, 221–225 (2012).
2. Qiu, W. *et al.* *Nature Struct. Mol. Biol.* <http://dx.doi.org/10.1038/nsmb.2205> (2012).
3. Vallee, R. B., Williams, J. C., Varma, D. & Barnhart, L. E. *J. Neurobiol.* **58**, 189–200 (2004).
4. Sakakibara, H. & Oiwa, K. *FEBS J.* **278**, 2964–2979 (2011).
5. Toba, S., Watanabe T. M., Yamaguchi-Okimoto, L., Toyoshima, Y. Y. & Higuchi, H. *Proc. Natl Acad. Sci. USA* **103**, 5741–5745 (2006).
6. Gennerich, A., Carter, A. P., Reck-Peterson, S. L. & Vale, R. D. *Cell* **131**, 952–965 (2007).
7. Warshaw, D. M. *et al.* *Biophys. J.* **88**, L30–L32 (2005).
8. Churchman, L. S., Ökten, Z., Rock, R. S., Dawson, J. F. & Spudis, J. A. *Proc. Natl Acad. Sci. USA* **102**, 1419–1423 (2005).
9. Rosenfeld, S. S., Fordyce, P. M., Jefferson, G. M., King, P. H. & Block, S. M. *J. Biol. Chem.* **278**, 18550–18556 (2003).
10. Ori-McKenney, K. M., Xu, J., Gross, S. P. & Vallee, R. B. *Nature Cell Biol.* **12**, 1228–1234 (2010).

PRECISION MEASUREMENT

A comb in the extreme ultraviolet

A 'comb' of photons at evenly spaced frequencies in the extreme ultraviolet has been generated. It will allow a more precise search for variation in the fine-structure constant, which sets the strength of the electromagnetic force. SEE LETTER P.68

LINDA YOUNG

That improved measurement precision yields new physics is axiomatic. In the case of optical spectroscopy, the quest for ultimate precision was revolutionized around the turn of the century by the realization of the optical frequency comb^{1–3}. This elegant invention, which was honoured by the 2005 Nobel Prize in Physics^{4,5}, provides a direct link between radio and optical frequencies, allowing one to count cycles of an electromagnetic field at near-petahertz frequencies (1 petahertz is 10¹⁵ Hz). On page 68 of this issue, Cingöz *et al.*⁶ extend the frequency range of combs to the extreme-ultraviolet spectral regime.

An optical frequency comb is produced by

a mode-locked ultra-fast laser in which a light pulse of sub-picosecond duration (1 picosecond is 10^{–12} s) circulates inside an optical cavity made up of a set of highly reflective mirrors. The laser emits an infinite train of pulses at a certain repetition rate (f_{rep}) that is determined by the time the pulses take to make a round trip of the cavity's mirrors^{7,8}. This train of pulses has two distinguishing properties: the pulses are equally spaced in time, and the phase (where a wave's peaks and troughs lie) of the electromagnetic field in each pulse is precisely shifted from that in the subsequent pulse (the locked-phase relationship). The train produces a comb of sharp spectral lines uniformly spaced in frequency.

The frequencies of the comb teeth are defined by a simple formula: $f_n = nf_{\text{rep}} + f_0$,

where f_{rep} and f_0 are radio frequencies, f_n is optical frequency and n is an integer of the order of 10⁶–10⁷ (Fig. 1). With one comb tooth, f_n , referenced to an absolute standard such as the frequency of a caesium atomic clock, the absolute frequency of an unknown optical laser may be counted as a beat note relative to another comb tooth with an accuracy at the level of one part in 10¹⁴. This greatly exceeds the accuracy achievable using conventional wavelength-based measurements, which are limited by wavefront distortion to an accuracy at the 10^{–10} level.

A prominent early example of the application of the frequency comb was the measurement¹ of a two-photon transition in hydrogen to an accuracy of more than 2 × 10^{–14}. This measurement, combined with one performed about four years later⁹, placed a limit on the variability of the fine-structure constant — a fundamental constant that characterizes the strength of the electromagnetic interaction and whose variation in time informs the standard model of particle physics. Beyond precision spectroscopy, the optical frequency comb has fostered a multitude of applications, including a clock based on optical frequencies, massively parallel multi-wavelength spectroscopy of unknown chemicals, and long-distance transmission of ultra-stable clock signals, which are useful for next-generation telescope arrays or accelerator systems.

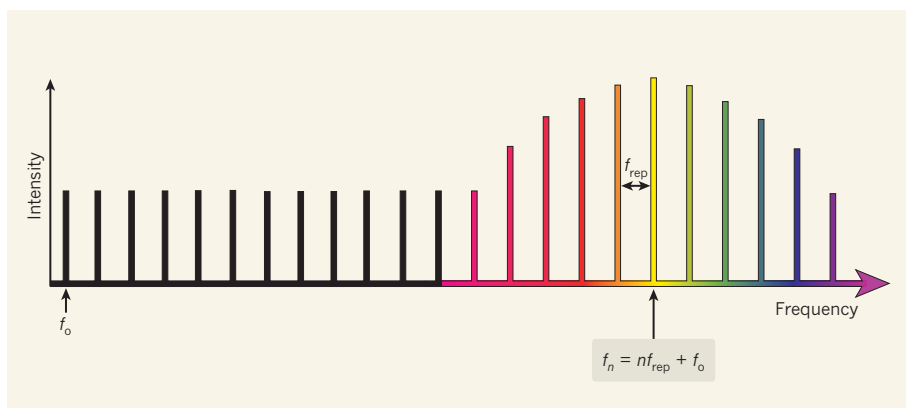


Figure 1 | Frequency comb. The intensity spectrum of a frequency comb consists of an array of sharp lines uniformly spaced in frequency by f_{rep} . The frequencies of the lines, f_n , are determined by a simple relationship, $f_n = n f_{\text{rep}} + f_0$, where f_0 is an offset frequency and n is an integer. Cingöz *et al.*⁶ demonstrate a frequency comb in the extreme-ultraviolet frequency range.

A question that naturally arises is whether frequency-comb technology can be extended to wavelengths shorter than those of the optical spectral range: to the ultraviolet or even the X-ray regime. A route to generating short-wavelength frequency combs that is directly linked to optical frequencies was pioneered by the ultra-fast-laser community, namely by researchers who study high-harmonic generation (HHG). HHG is the process of generating light of frequencies that are multiples (high harmonics) of a fundamental frequency, and is achieved by focusing intense optical pulses into a gaseous target.

In the case of a single atom, HHG involves three steps: ionization of the atom by a laser field; propagation of the resulting electron in the laser field; and recombination of the electron with the parent ion in a process that is accompanied by the emission of high-harmonic light. Although this model captures most of the underlying physics, a complete picture requires an understanding of how waves propagate through partially ionized media¹⁰. Because the laser intensity required for this extreme nonlinear process to occur is high (more than 10^{13} watts per square centimetre), amplified femtosecond lasers at low repetition rates (10 Hz to 100 kHz) with 10^{-3} joules per pulse are typically used. The 10^{-8} -joule pulses of a typical 100-MHz optical frequency comb are simply not powerful enough for HHG.

In 2005, researchers reported^{11,12} that the requisite intensity for HHG could be obtained by adding pulses from an optical frequency comb using an ‘enhancement’ or ‘power build-up’ cavity, in which successive pulses add constructively because the phases of the pulses’ electromagnetic waves are locked. In such a cavity, the pulse energy can be multiplied nearly 1,000-fold. In addition, it is essential for the HHG process to retain the locked-phase relationship over many consecutive pulses to preserve the underlying comb structure (about 10^5 teeth spaced by f_{rep}

for each harmonic) and effectively make the resulting HHG a shorter-wavelength comb. In these two early experiments^{11,12}, the frequency comb originating from a mode-locked titanium-sapphire laser (with a power of 1 W, repetition rate of about 100 MHz and pulses of roughly 50 fs) was fed into an enhancement cavity containing a jet of xenon gas located at an intracavity focus to produce HHG, in one case¹¹ up to the 15th harmonic.

However, simply detecting HHG radiation is not sufficient to verify that comb structure exists. Therefore, both groups^{11,12} mixed third-harmonic radiation at a wavelength of 266 nm with an independently derived reference signal to observe beat signals that have a spectral linewidth of a few hertz, much less than the roughly 10^{12} Hz expected from the uncertainty principle for 50-fs pulses. These demonstrations were a solid first step towards realizing extreme-ultraviolet combs. But the extracted power per comb tooth was insufficient for applications such as precision spectroscopy, and comb structure was verified up to only the third harmonic.

In their study, Cingöz *et al.*⁶ lift these barriers to a more widespread application of extreme-ultraviolet-frequency combs. In contrast to the early demonstrations, the authors show comb structure out to the 13th harmonic (82 nm), and with enough power (200 μ W per harmonic) for applications. They achieved these results by using the highest-power near-infrared frequency comb, which is based on an 80-W, 120-fs ytterbium-fibre laser¹³ coupled into an enhancement cavity of novel design. The design strategy was informed by a better understanding of the intracavity HHG process¹⁴. Nonlinearities in the HHG process at the intracavity focus clamp the peak power of the circulating pulses and lead to instabilities that make the constructive addition of the pulses difficult¹⁵. Cingöz *et al.* used an enhancement cavity that is more robust but loses more power. In this cavity, pulses are amplified only 200-fold, compared with the

original 600-fold used by the same group¹².

Cingöz and colleagues went on to perform single-photon spectroscopy of an argon transition using an individual comb tooth (with a linewidth of 11 MHz) with an unprecedentedly high accuracy for the extreme-ultraviolet spectral region. They repeated this measurement as a function of f_{rep} , unambiguously establishing the comb structure at the 13th harmonic. However, single-photon spectroscopy is limited to probing narrow, well-isolated absorption lines; here it was performed with a single comb tooth containing about 10 pW of power. Notably, spectroscopy of two-photon transitions with a frequency comb can use a much larger fraction of the total comb power because pairs of comb teeth whose frequencies sum to the transition frequency also drive the transition.

The authors’ demonstration of an extreme-ultraviolet frequency comb, which allows robust linkage between the radio and extreme-ultraviolet spectral regions, ushers in an era of precision metrology in the extreme ultraviolet. A recognized application of their extreme-ultraviolet comb is the aforementioned search for variability in the fine-structure constant: transitions in highly charged ions that have energy in the extreme ultraviolet range are calculated¹⁶ to have higher sensitivity to variations in the fine-structure constant than do optical transitions. As powerful extreme-ultraviolet frequency combs^{6,17} are developed that enable a higher precision than that achieved by the optical frequency comb, it will be exciting to see what other applications might emerge. ■

Linda Young is in the X-ray Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA.
e-mail: young@anl.gov

- Reichert, J. *et al.* *Phys. Rev. Lett.* **84**, 3232–3235 (2000).
- Jones, D. J. *et al.* *Science* **288**, 635–639 (2000).
- Diddams, S. A. *et al.* *Phys. Rev. Lett.* **84**, 5102–5105 (2000).
- Hall, J. L. *Rev. Mod. Phys.* **78**, 1279–1295 (2006).
- Hänsch, T. W. *Rev. Mod. Phys.* **78**, 1297–1309 (2006).
- Cingöz, A. *et al.* *Nature* **482**, 68–71 (2012).
- Udem, Th., Holzwarth, R. & Hänsch, T. W. *Nature* **416**, 233–237 (2002).
- Cundiff, S. T. & Ye, J. *Rev. Mod. Phys.* **75**, 325–342 (2003).
- Fischer, M. *et al.* *Phys. Rev. Lett.* **92**, 230802 (2004).
- Brabec, T. & Krausz, F. *Rev. Mod. Phys.* **72**, 545–591 (2000).
- Gohle, C. *et al.* *Nature* **436**, 234–237 (2005).
- Jones, R. J., Moll, K. D., Thorpe, M. J. & Ye, J. *Phys. Rev. Lett.* **94**, 193201 (2005).
- Ruehl, A., Marcinkewicz, A., Fermann, M. E. & Hartl, I. *Opt. Lett.* **35**, 3015–3017 (2010).
- Allison, T. K., Cingöz, A., Yost, D. C. & Ye, J. *Phys. Rev. Lett.* **107**, 183903 (2011).
- Yost, D. C. *et al.* *Opt. Exp.* **19**, 23483–23493 (2011).
- Berengut, J. C., Dzuba, V. A., Flambaum, V. V. & Ong, A. *Phys. Rev. Lett.* **106**, 210802 (2011).
- Lee, J., Carlson, D. R. & Jones, R. J. *Opt. Exp.* **19**, 23315–23326 (2011).

Conditional modulation of spike–timing–dependent plasticity for olfactory learning

Stijn Cassenaer^{1,2} & Gilles Laurent^{1,3}

Mushroom bodies are a well-known site for associative learning in insects. Yet the precise mechanisms that underlie plasticity there and ensure their specificity remain elusive. In locusts, the synapses between the intrinsic mushroom body neurons and their postsynaptic targets obey a Hebbian spike–timing–dependent plasticity (STDP) rule. Although this property homeostatically regulates the timing of mushroom body output, its potential role in associative learning is unknown. Here we show *in vivo* that pre–post pairing causing STDP can, when followed by the local delivery of a reinforcement–mediating neuromodulator, specify the synapses that will undergo an associative change. At these synapses, and there only, the change is a transformation of the STDP rule itself. These results illustrate the multiple actions of STDP, including a role in associative learning, despite potential temporal dissociation between the pairings that specify synaptic modification and the delivery of reinforcement–mediating neuromodulator signals.

Behavioural and genetic experiments in *Drosophila* and honeybees have revealed that the mushroom body, a brain area containing up to hundreds of thousands of neurons called Kenyon cells, is critical for associative learning of odours^{1–11} but not for the expression of innate odour-driven behaviours⁵ (Fig. 1a). Recent electrophysiological experiments in locusts and other insects show that the responses of Kenyon cells to odours are highly selective and, thus, rare^{12–14}. By contrast, antennal lobe neurons, the source of the olfactory input to Kenyon cells, are few and promiscuous¹². Odour codes are thus ‘compact’ in the antennal lobe—that is, the representation of each odour engages many neurons in a small population—but ‘sparse’ in the mushroom bodies (Fig. 1a). Although sparse codes require larger neuron populations, they are beneficial for memory because they can reduce interference between memory traces^{12,15,16}. Kenyon cells project to two regions, called α - and β -lobes, where they synapse onto small populations of ‘extrinsic’ neurons. In locusts, the synapses between Kenyon cells and β -lobe neurons (bLNs) are modifiable by a Hebbian STDP rule^{17–19} (Fig. 1a), but nothing so far implicates STDP in associative learning there; rather, STDP causes the homeostatic regulation of bLN spike timing¹⁷. Recent experimental results in moths¹⁴ show that Kenyon cell responses to odours recorded during behavioural learning generally occur and end well before reward delivery, indicating that STDP alone cannot support associative conditioning^{20,21}. Neuromodulation has been proposed recently as a potential solution^{22,23}. We address this issue with *in vivo* electrophysiology in locusts and discover a complex interplay between STDP, reinforcer signals and odour codes in mushroom bodies.

Dense odour representations in β -lobes

We first examine odour representations in the β -lobes. Using intradendritic recordings, we sampled the responses of 55 bLNs to up to 16 odours (Methods and Fig. 1b). Each bLN responded to nearly every odour, with responses that differed in intensity, patterning, delay and duration across neuron–odour pairs. The probability that a given bLN responded to an odour was 0.97. On average, a bLN fired action potentials in approximately half of the local field potential (LFP) oscillation cycles during the odour presentation (0.502 ± 0.219 , $n = 126$

neuron–odour pairs, 25 LFP cycles per response): on average, half of the population was active in any given cycle. Hence, these representations resemble those in the antennal lobes¹² (and exceed them in promiscuity) rather than the sparse representations by Kenyon cells, to which bLNs are directly connected.

We tested whether the broad tuning of bLNs might be explained by known features of mushroom body circuits. We implemented a simple model (Methods and ref. 17) constrained by Kenyon cell response statistics and timing¹², by the properties of STDP at Kenyon cell synapses¹⁷ and by Kenyon cell/bLN (KC–bLN) connectivity ratios estimated from experiments¹⁷. With such a model, we could reproduce the bLN firing phase observed experimentally¹⁷ (Fig. 2a, left) but not the odour-response intensity or probability (Fig. 2a, right, and Supplementary Fig. 1.1a). Rather, activity across the model bLN (mbLN) population saturated rapidly when STDP was turned on. This behaviour is a known property of rate-based Hebbian learning in model networks and can be counteracted by imposing synaptic weight bounds or renormalization rules²⁴.

Inhibition limits STDP and saturation

In examining our experimental data on bLN responses to odours, we observed that the periodic excitatory input originating from Kenyon cells was often curtailed by phasic inhibitory postsynaptic potentials (Fig. 2b, arrowheads), with onsets at the typical phase of bLN action potentials. We hypothesized that these inhibitory postsynaptic potentials originate from lateral inhibition among bLNs (as put forward in a theoretical exploration²⁵). This was supported by extracellular stimulation of Kenyon cells (Supplementary Fig. 2) and confirmed in paired bLN recordings. Beta-lobe neurons inhibit each other (Fig. 2c) with an estimated connection probability of 28% ($n = 32$ connections, 2 reciprocally connected pairs). Although unknown so far, other interneuron populations may also contribute to bLN inhibition.

In each oscillation cycle, lateral inhibition reduces the likelihood of late bLN spikes; thus, it should limit the ability of STDP to potentiate KC–bLN synapses. This in turn should curb bLN population activity (see also ref. 26). As predicted, implementing lateral inhibition

¹Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. ²Broad Fellows Program in Brain Circuitry, California Institute of Technology, Pasadena, California 91125, USA.

³Max Planck Institute for Brain Research, 60528 Frankfurt am Main, Germany.

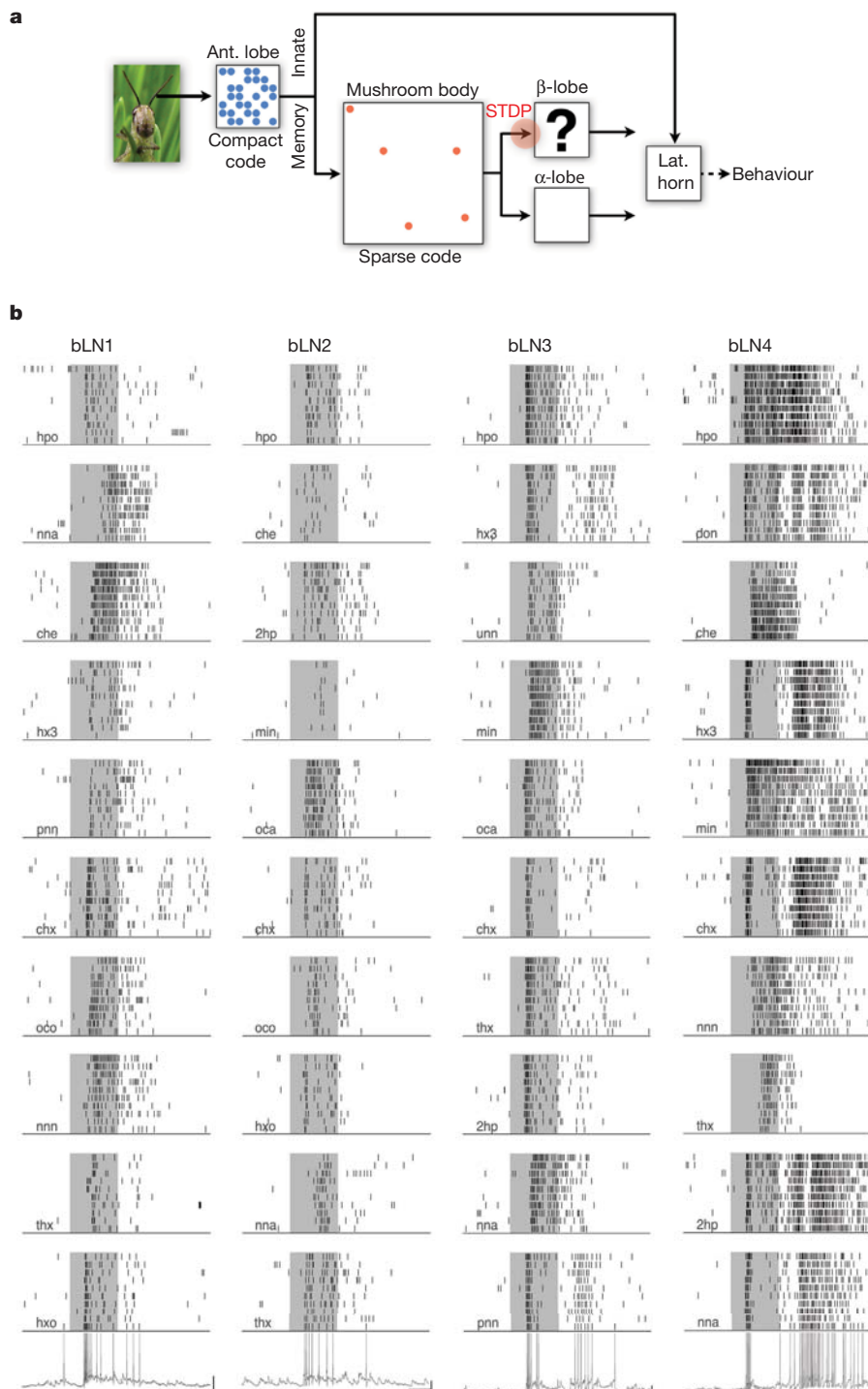


Figure 1 | Beta-lobe neurons are promiscuous. **a**, Olfactory circuit in which bLNs are embedded. Ant., antennal; lat., lateral. **b**, Rasters constructed from *in vivo* intracellular dendritic recordings, illustrating responses of four bLNs (different animals) to ten odours. Shaded region represents odour delivery (1 s). For each bLN, intracellular traces at bottom correspond to the tenth trial with the last odour.

between mbLNs (based on the experimental connectivity estimate and amplitude distribution) reduced population activity to $\sim 50\%$ ($p_{\text{response}} = 0.51 \pm 0.14$ (response probability), $n = 30$ mbLNs; compare with $p_{\text{response}} = 0.49 \pm 0.18$, $n = 30$ bLNs from experiment; Fig. 2d–f). The mBLN spike phase remained centred on π (LFP trough). In addition, the response profiles of those mbLNs were very similar to those recorded experimentally (compare distributions in Fig. 2e, f and moments in Fig. 2g; see also Supplementary Figs 3–5).

Horizontal scale bars, 500 ms; vertical scale bars, 10 mV; spikes are clipped. Odours: 2-heptanone (2hp), cherry (che), *cis*-3-hexen-1-ol (chx), 3,7-dimethyl-2,6-octadiene-nitrile (don), 1-heptanol (hpo), 1-hexen-3-ol (hx3), 1-hexanol (hxo), mint (min), nonanal (nna), 5-nonanone (nnn), octanal (oca), 1-octanol (oco), 3-pentanone (pnn), *trans*-2-hexen-1-ol (thx), 6-undecanone (unn).

We conclude that bLN activity, although very different from the sparse output of Kenyon cells, can be explained by the combined effects of Hebbian STDP at KC–bLN synapses and lateral inhibition between bLNs. The useful consequences are that bLN spikes are tightly phase-locked to the LFP¹⁷ and that instantaneous bLN population activity lies around the midpoint of its dynamic range. This increases the coding capacity of this small population of output neurons and is relevant for plasticity, as explored below.

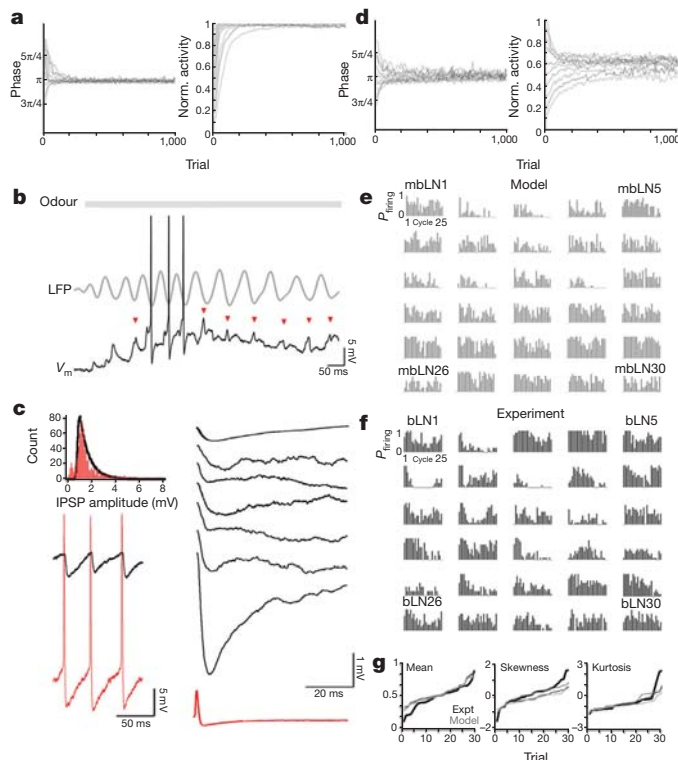


Figure 2 | Saturation of activity caused by STDP can be counteracted by lateral inhibition. **a**, Mean firing phase (left) and normalized population activity (right) in a simulation of a KC–bLN network (30 Izhikevich units⁴⁴, one LFP cycle, 1,000 trials). Each point is a ten-trial average; STDP is turned on in trial 11. Each curve is run with n Kenyon cell inputs per mbLN ($20 < n < 80$, 10% connections in common). STDP regulates bLN firing phase (ref. 17) but inexorably drives bLN population to saturation. **b**, *In vivo* intradendritic recording of bLN membrane potential, V_m , together with LFP during odour presentation. Arrowheads indicate onsets of inhibitory postsynaptic potentials, timed on LFP troughs. **c**, Top left: frequency distribution of inhibitory postsynaptic potential amplitudes, measured at baseline, in 30 bLNs; black curve, fit used for model in **d**. IPSP, inhibitory postsynaptic potential. Bottom left: paired bLN recording revealing direct inhibitory connection. Right: bLN spike-triggered averages from seven distinct bLN pairs. **d**, Simulation of a bLN network as in **a**, except with n Kenyon cell inputs per bLN (here $22 < n < 130$, same drive at simulation onset as in **a**). Lateral inhibition among bLNs is as determined experimentally. Mean phase (left) is minimally affected; network activity (right) settles near midrange. **e**, Simulation of a bLN network as in **d**, extended over 25 LFP cycles, with time-varying Kenyon cell input profiles derived from experiments^{12,17,45}. Post-stimulus time histograms plotted as firing probability versus cycle rank for each of the 30 mbLNs (ten trials), post STDP. **f**, *In vivo* intradendritic responses of 30 bLNs to *trans*-2-hexen-1-ol represented as in **e** (from 16 different animals). **g**, Comparison of experimental (black) and model (grey) bLN response statistics. Same simulation as in **e** (darkest grey) together with two other simulations of the same network with different Kenyon cell input profiles (lighter shades of grey).

Conditional modulation of STDP rule

Mushroom bodies are critical for associative conditioning in insects^{6–11,27}, and second-messenger pathways involved in memorization and retrieval have been described in *Drosophila*^{1,4–7,28,29}. Yet we know little about the synaptic mechanisms underlying conditioning there³⁰, and even less about how specificity is achieved or about the roles that STDP might have in it. Associative learning in insects seems to rely on neuromodulators^{31–35} (for example octopamine (OCT) and dopamine) to mediate unconditioned signals (appetitive and aversive, respectively, for OCT and dopamine). In locusts and other insects, these neuromodulators are delivered nonspecifically by small numbers of octopaminergic³⁶ and dopaminergic³⁷ neurons with projections to the mushroom body lobes (among others), that is, where Kenyon cells contact bLNs.

We assessed the effect of local OCT injection into the β -lobes on plasticity and tuning of those bLNs *in vivo* (Fig. 3). Figure 3d illustrates our experimental schemes, which were designed to probe specificity, that is, the question of how the system determines which synapses should change when the reinforcer is global. During an odour stimulus, two features distinguish odour-evoked Kenyon cell spikes from ‘stray’ spikes produced randomly by non-responding Kenyon cells (most of the Kenyon cells): odour-evoked spikes are synchronized and cause bLNs to spike by their collective action. Hence, odour-evoked release at Kenyon cell synapses occurs within the time window for STDP. We thus tested the effect of OCT injection on synapses that differed only in their temporal features of pre- and postsynaptic activation. In each experiment, we measured the responses of one bLN to Kenyon cell stimulation at a rate of 0.1 s^{-1} at two locations (a pairing location (P) and a control location (C)) and two different times (C 300 ms before P). We found Kenyon cell stimulation sites that converged onto the same recorded bLN in 13 out of 20 experiments.

In ‘test’ trials (Fig. 3d, top), the bLN was depolarized by a brief pulse of current to produce one single action potential 5 s after Kenyon cell stimulation at P, hence well outside the window of $\delta t = t_{\text{post}} - t_{\text{pre}} = \pm 25 \text{ ms}$ required for STDP¹⁷, where t_{post} and t_{pre} are respectively the post- and presynaptic spike times. Pairing consisted of five trials at 0.1 s^{-1} , of either of two kinds. In ‘STDP’ trials, the bLN was fired within $\pm 25 \text{ ms}$ of a Kenyon cell stimulus at P (Fig. 3d, middle). In ‘STDP + OCT’ trials, pairing was as in STDP trials but was followed by one OCT injection (5 pl) 1 s after pairing (Fig. 3d, bottom). Test trials were carried out before STDP trials (Fig. 3a–c, red), after STDP trials (black) and after STDP + OCT trials (blue). All test and pairing trials thus contained the same numbers of pre- and postsynaptic pulses; only their relative timings differed. Examples with three values of δt are shown (Fig. 3a–c). In each, OCT caused a depression. We note that only connections between cells co-activated within the STDP window were depressed (that is, P but not C; see also Supplementary Fig. 6). The depressive action of OCT on synapses undergoing STDP (20 different values of δt ; Fig. 3e, f and Supplementary Fig. 7) was greatest for $\delta t > 0$, converting baseline Hebbian STDP into one where only depression occurred.

To rule out the possibility that the action of OCT resulted not from the pairing *per se*, but from the temporal proximity of OCT delivery to the pairing (P), we reversed the order of C and P ($n = 3$; Supplementary Fig. 8). Again, only the P pathway was affected by OCT (effect of OCT on control (mean \pm s.e.m.): $-0.9 \pm 0.49\%$, $n = 3$ (P then C); $0.2 \pm 2.28\%$, $n = 10$ (C then P); sets statistically equivalent: $P > 0.66$, t -test). Hence, because the control and paired synapses shared the same postsynaptic partner (C and P differed only in the identity of the presynaptic cells and in the timings of the presynaptic spikes relative to those in their common bLN), and because the two sets of presynaptic cells were stimulated within only 300 ms of one another (C then P or P then C), we conclude that the relative timing of pre- and postsynaptic firing is what contributes to making the synapses susceptible to OCT action. We do not know whether pre-post delays longer than $\pm 25 \text{ ms}$ (but shorter than 300 ms) would have equivalent effects. However, almost all odour-evoked Kenyon cell action potentials occur within the range of delays we tested. We conclude that, despite its global release, OCT affects specific synapses—those in which pre- and postsynaptic partners fired within the appropriate delay between one another. This is the case even though OCT was delivered 1 s after the relevant pairing (reproducing a temporal dissociation between conditioned stimulus and unconditioned stimulus¹⁴).

Conditional changes of STDP are specific

The STDP curves in Fig. 3f were obtained *in vivo* with electrical activation of the pre- and postsynaptic cells. We next tested the effect of OCT on the responses of bLNs to odours. We predicted that if OCT depresses KC–bLN synapses that fire within the STDP time window,

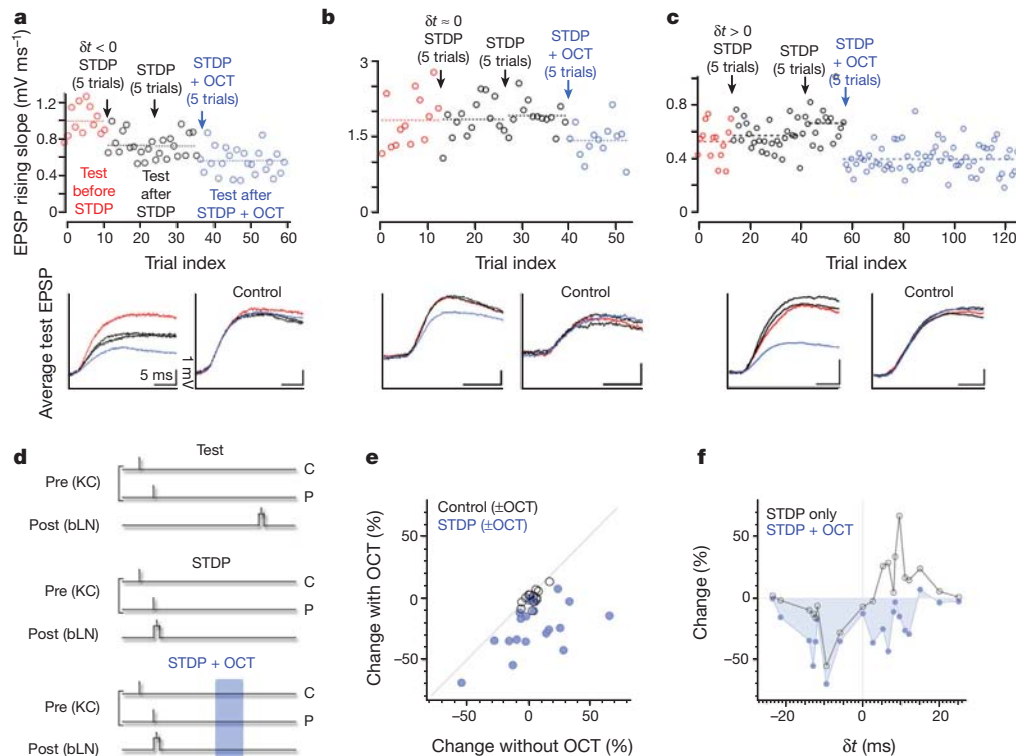


Figure 3 | OCT changes the STDP rule. **a**, Five trials pairing OCT with depressing STDP ($\delta t < 0$) reduced KC–bLN excitatory postsynaptic potential (EPSP) size (blue) more than did two sets of five-trial STDP with negative δt alone (black). EPSPs evoked by interleaved stimulation with the control electrode (embedded among Kenyon cells in a different location) were unaffected (control, bottom right), despite equal exposure to OCT. **b**, Five trials pairing OCT with $\delta t \approx 0$ STDP reduced KC–bLN EPSP size (blue), whereas STDP alone with this δt value had no effect (black). Control EPSPs were unaffected (bottom right). **c**, Five trials pairing OCT with potentiating STDP ($\delta t > 0$) reduced KC–bLN EPSP size after it was potentiated by two bouts of STDP alone with this δt value. Control EPSPs were unaffected (bottom right). **d**, Stimulation protocols. Test trials were used to evaluate the amplitudes of KC–bLN EPSPs before STDP, after STDP and after STDP + OCT, as shown in **a–c**, and to compute (from multiple experiments) values in **e** and **f**. **c**, control Kenyon cell input; **P**, paired Kenyon cell input. In test trials, bLNs were depolarized with a 5-ms d.c. pulse (causing a single bLN spike) ~ 5 s after

extracellular stimulation (0.1-ms pulses) of Kenyon cells in two different locations (C and P, 300 ms apart). In STDP trials, bLNs were depolarized as above, causing a single spike in a narrow window around the time of Kenyon cell stimulation (P). Control Kenyon cell stimulation (C) was offset by 300 ms. STDP + OCT trials were the same as STDP trials, only with the addition of localized OCT injection (50-ms pressure-pulse) in the β -lobe, 1 s after P Kenyon cell stimulation. **e**, Summary data from 20 experiments, plotting for each experiment the effect of STDP + OCT versus STDP alone (blue dots) and plotting control data, with and without OCT, for 13 of 20 experiments (black circles). Changes due to STDP + OCT differ significantly from control + OCT ($P < 0.003$, t -test). Changes due to OCT in control not significantly different from zero ($P > 0.84$, t -test; see also Supplementary Fig. 7). **f**, Comparison between change due to STDP + OCT and change due to STDP alone, each plotted as a function of δt . Changes due to STDP + OCT differ significantly from zero (relative to before STDP: $P < 0.0001$; relative to STDP only: $P < 0.0002$, t -test; see also Supplementary Fig. 7).

it should also reduce bLN output. This was tested first with our model, using an STDP rule modified according to Fig. 3f: responses were reduced in 28 of 30 mBLNs (Fig. 4a, b). We then tested the prediction experimentally with 12 bLNs (18 bLN–odour pairs), each recorded intracellularly from a dendrite. In 16 of 18 instances, OCT caused a reduction of bLN output, averaged over the odour response (Fig. 4c, d). LFP power in the mushroom body calyx was unaffected (Fig. 4f), consistent with unchanged input to the mushroom body.

We also evaluated the specificity of changes in bLN–odour responses. A bLN was presented with three odours in blocks of trials in the order indicated (Fig. 4e, left). OCT was injected 800 ms after odour onset, paired each time with only one of the three odours. As in Fig. 4c, d, the responses of the bLN to the three odours (assembled by odour in three graphs; Fig. 4e, right) were reduced after pairing with OCT (compare blue and red post-stimulus time histograms). The reduction of the response to an odour occurred only after pairing of that particular odour with OCT. For example, responses to cherry and *trans*-2-hexen-1-ol were minimally affected by pairing OCT with 2-heptanone; they became significantly reduced only after the appropriate pairing of either cherry or *trans*-2-hexen-1-ol with OCT. Figure 4f compares summary data for odours paired with OCT ($n = 18$ bLN–odour pairs) to controls ($n = 12$ bLN–odour pairs:

odour responses after pairing of a different odour with OCT ($n = 7$) and odour responses after pairing of electrical Kenyon cell stimulation with OCT ($n = 5$); $n = 18$ LFP recordings simultaneous with intracellular bLN recordings). These results support the observation that OCT specifically depresses contacts between cells co-activated within the STDP window (Fig. 3); they also indicate that KC–bLN pairing and OCT should co-occur within an interval of at most a few tens of seconds.

Discussion

In the absence of OCT, STDP at KC–bLN synapses adaptively regulates the bLN spike phase¹⁷ and bLN population output, helped by the action of lateral inhibition between bLNs (Fig. 2). In the context of reinforcer delivery, STDP gains a new role: it tags—presumably through biochemical changes dependent on pairing—the synapses between connected Kenyon cells and bLNs that have fired an action potential within the STDP pairing window, itself commensurate with the oscillation period in this system³⁸. This renders those synapses susceptible to the action of OCT (which is proposed to mediate positive reinforcement in insect learning^{14,34,39}) and reduces the inhibitory output of the mushroom body mediated by bLNs. Hence, the action of OCT can be specific, even though its release is diffuse and often delayed

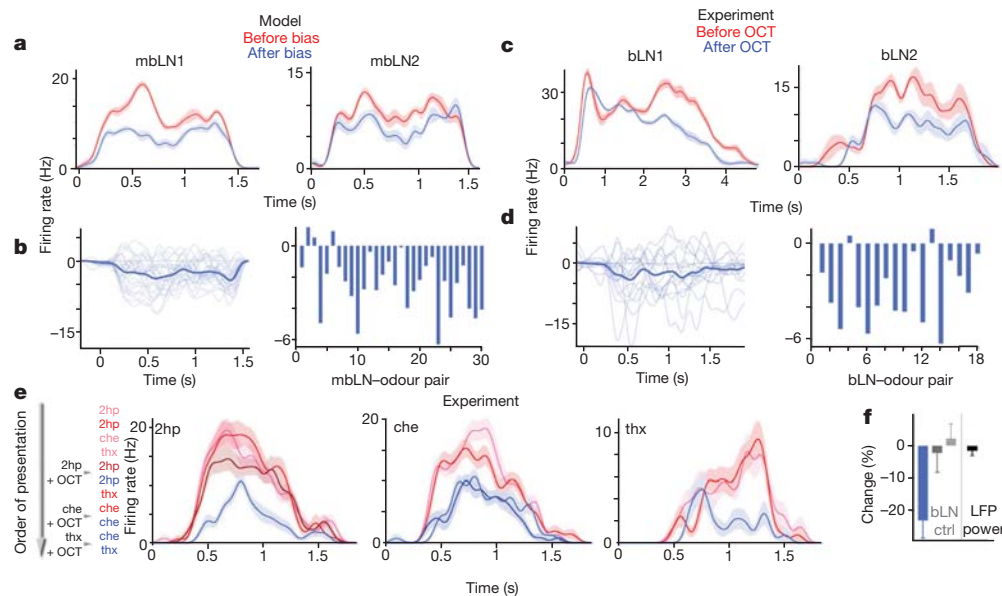


Figure 4 | Odour-specific decrease of mushroom body output by OCT.

a, Simulation of bLN network as in Fig. 2e. Smoothed post-stimulus time histograms (PSTHs) for two mblNs (ten trials; shading, s.e.). Red, responses at equilibrium; blue, responses after ten trials with STDP + OCT curve shown in Fig. 3f. **b**, Left: difference in PSTHs for all 30 mblNs responses before and after five trials of STDP + OCT ('PSTH after pairing' minus 'PSTH before pairing', thin lines). Responses of individual units at particular times can increase or decrease, but average difference (thick line) is less than zero at all times. Right: time-averaged differences for each of the 30 mblNs ('PSTH after pairing' minus 'PSTH before pairing'). In some instances, mblN responses increased (for most units this occurred transiently; for a few mblNs the entire time-averaged response increased). This is due to reduced lateral inhibition (that is, disinhibition) because mblN population activity decreased. **c**, *In vivo* intradendritic recordings of bLN odour responses before and after pairing of the odour with localized OCT injection in the β -lobe (50-ms pressure-pulse, 800 ms after odour onset, either five ($n = 5$) or ten pairing trials ($n = 13$)). PSTHs for two bLNs (ten trials; shading, s.e.): red, responses before odour–OCT pairing; blue, responses after ten trials of odour paired with OCT injection. **d**, Same difference of PSTHs as in **b**, for 18 recorded bLN–odour pairs (12 bLNs). As in the model, responses of individual bLNs can increase or

decrease transiently, but on average there is a net decrease. **e**, OCT-induced changes are odour specific. Three odours are sequentially paired (during five trials) with OCT to assess carry-over of pairing across odours. Order of stimuli is shown at left. Red colours, responses before pairing with OCT; blue colours, responses after pairing. Odours are presented in ten-trial blocks and PSTHs are shown for each block (shading, s.e.). Changes due to nonspecific pairing are minimal by comparison with those due to odour-specific pairing. **f**, Summary data comparing the effect of odour–OCT pairing (blue: $n = 18$ bLN–odour pairs, 12 bLNs) with those of controls (grey: $n = 12$ bLN–odour pairs, 9 bLNs; black (LFP controls): $n = 18$ recordings simultaneous with intracellular bLN recordings). Dark grey ($n = 7$), effect on bLN response to non-paired odour before and after pairing another odour with OCT. Light grey ($n = 5$), effect on bLN response to non-paired odour before and after pairing electrical activation of Kenyon cells with OCT. Black, effect on LFP power (10–30-Hz band, recorded in mushroom body) before and after pairing odour with OCT. Changes due to pairing odour with OCT differ significantly from zero ($P < 0.0006$, t -test) and from those in the three control conditions ($P < 0.002$ – 0.022 , t -test). Changes in control conditions were not significantly different from zero ($P > 0.72$ (dark grey), 0.64 (light grey) and 0.33 (black)).

relative to the conditioned (odour) stimulus. This provides experimental support for related theoretical propositions^{23,40,41}. Our results also solve the temporal dissociation puzzle revealed in experiments¹⁴.

Because the selectivity of synaptic change allowed by STDP relies on the rare co-occurrence of pre- and postsynaptic spikes, it is beneficial that Kenyon cell firing is rare and that odour representations by Kenyon cells are sparse, as observed^{12–14}. We propose that olfactory associative memories are stored, at least in part, as sparse sets of synaptic weights between Kenyon cells and mushroom body output neurons. In the absence of reinforcer, STDP and lateral inhibition between bLNs set the (default) instantaneous output of the β -lobe network near the middle of its dynamic range. This serves at least two purposes. First, it maximizes the number of possible distinct combinations of simultaneously active bLNs. Second, it ensures that the output of the network can be increased or decreased, if an associative reinforcer is to co-occur. It is possible that some of the molecular pathways revealed by *Drosophila* neurogenetics^{6,9,11} have a role in the phenomenon described here. The functional form of STDP has recently been shown to depend on neuromodulatory context in mammalian brain slices^{22,42,43}, suggesting that our results may have broad implications for STDP and learning.

METHODS SUMMARY

All results were obtained *in vivo* from locusts housed in a crowded colony. Odours were delivered by injection of a controlled volume of odorized air within a

constant stream of desiccated air. The results presented here are derived from intradendritic recordings, made using sharp micropipettes, of 76 neurons in 55 locusts. LFPs were recorded with silicon probes (NeuroNexus) in the mushroom body Kenyon cell soma cluster. Octopamine (100 μ M) was injected locally into the mushroom body β -lobe using a picopump (WPI). Electrical stimulation of Kenyon cell somata was done with modified tetrodes (FHC). Stimulation protocols and data analysis were carried out using specialized software (LABVIEW, National Instruments; IGOR, Wavemetrics; and MC_STIMULUS, Multichannel Systems). Simulations of networks of 'Izhikevich units'⁴⁴ were done using MATLAB (Mathworks).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 September; accepted 8 December 2011.

Published online 25 January 2012.

- McGuire, S. E., Le, P. T. & Davis, R. L. The role of *Drosophila* mushroom body signaling in olfactory memory. *Science* **293**, 1330–1333 (2001).
- Tully, J. B. C. T. in *Drosophila: A Practical Approach* (ed. Roberts, D. B.) 265–317 (Oxford Univ. Press, 1998).
- de Belle, J. S. & Heisenberg, M. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–695 (1994).
- Tully, T. & Quinn, W. G. Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. A* **157**, 263–277 (1985).
- Heisenberg, M., Borst, A., Wagner, S. & Byers, D. *Drosophila* mushroom body mutants are deficient in olfactory learning. *J. Neurogenet.* **2**, 1–30 (1985).
- Quinn, W. G., Harris, W. A. & Benzer, S. Conditioned behavior in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **71**, 708–712 (1974).

7. Keene, A. C. & Waddell, S. *Drosophila* olfactory memory: single genes to complex neural circuits. *Nature Rev. Neurosci.* **8**, 341–354 (2007).
8. Akalal, D. B. *et al.* Roles for *Drosophila* mushroom body neurons in olfactory learning and memory. *Learn. Mem.* **13**, 659–668 (2006).
9. Davis, R. L. Olfactory memory formation in *Drosophila*: from molecular to systems neuroscience. *Annu. Rev. Neurosci.* **28**, 275–302 (2005).
10. Gerber, B., Tanimoto, H. & Heisenberg, M. An engram found? Evaluating the evidence from fruit flies. *Curr. Opin. Neurobiol.* **14**, 737–744 (2004).
11. Heisenberg, M. Mushroom body memoir: from maps to models. *Nature Rev. Neurosci.* **4**, 266–275 (2003).
12. Perez-Orive, J. *et al.* Oscillations and sparsening of odor representations in the mushroom body. *Science* **297**, 359–365 (2002).
13. Turner, G. C., Bazhenov, M. & Laurent, G. Olfactory representations by *Drosophila* mushroom body neurons. *J. Neurophysiol.* **99**, 734–746 (2008).
14. Ito, I., Ong, R. C.-Y., Raman, B. & Stopfer, M. Sparse odor representation and olfactory learning. *Nature Neurosci.* **11**, 1177–1184 (2008).
15. Kanerva, P. *Sparse Distributed Memory* (MIT Press, 1988).
16. Papadopoulou, M., Cassenaer, S., Nowotny, T. & Laurent, G. Normalization for sparse encoding of odors by a wide-field interneuron. *Science* **332**, 721–725 (2011).
17. Cassenaer, S. & Laurent, G. Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature* **448**, 709–713 (2007).
18. Markram, H., Lübke, J., Frotscher, M. & Sakmann, B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**, 213–215 (1997).
19. Bi, G. Q. & Poo, M. M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10464–10472 (1998).
20. Abbott, L. F. & Nelson, S. B. Synaptic plasticity: taming the beast. *Nature Neurosci.* **3**, 1178–1183 (2000).
21. Meeks, J. P. & Holy, T. E. Pavlov's moth: olfactory learning and spike-timing-dependent plasticity. *Nature Neurosci.* **11**, 1126–1127 (2008).
22. Pawlak, V. Timing is not everything: neuromodulation opens the STDP gate. *Front. Syn. Neurosci.* **2**, 146 (2010).
23. Izhikevich, E. M. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* **17**, 2443–2452 (2007).
24. Gütig, R., Aharonov, R., Rotter, S. & Sompolinsky, H. Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J. Neurosci.* **23**, 3697–3714 (2003).
25. Huerta, R., Nowotny, T., García-Sánchez, M., Abarbanel, H. D. I. & Rabinovich, M. I. Learning classification in the olfactory system of insects. *Neural Comput.* **16**, 1601–1640 (2004).
26. Masquelier, T. & Thorpe, S. J. Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Comput. Biol.* **3**, e31 (2007).
27. Waddell, S. & Quinn, W. G. Learning how a fruit fly forgets. *Science* **293**, 1271–1272 (2001).
28. Isabel, G., Pascual, A. & Preat, T. Exclusive consolidated memory phases in *Drosophila*. *Science* **304**, 1024–1027 (2004).
29. Gervasi, N., Tchénio, P. & Preat, T. PKA dynamics in a *Drosophila* learning center: coincidence detection by rutabaga adenylyl cyclase and spatial regulation by dunce phosphodiesterase. *Neuron* **65**, 516–529 (2010).
30. Wang, Y., Mamiya, A., Chiang, A. S. & Zhong, Y. Imaging of an early memory trace in the *Drosophila* mushroom body. *J. Neurosci.* **28**, 4368–4376 (2008).
31. Claridge-Chang, A. *et al.* Writing memories with light-addressable reinforcement circuitry. *Cell* **139**, 405–415 (2009).
32. Kim, Y. C., Lee, H. G. & Han, K. A. D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in *Drosophila*. *J. Neurosci.* **27**, 7640–7647 (2007).
33. Riemensperger, T., Voller, T., Stock, P., Buchner, E. & Fiala, A. Punishment prediction by dopaminergic neurons in *Drosophila*. *Curr. Biol.* **15**, 1953–1960 (2005).
34. Schwaerzel, M. *et al.* Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in *Drosophila*. *J. Neurosci.* **23**, 10495–10502 (2003).
35. Mizunami, M. & Matsumoto, Y. Roles of aminergic neurons in formation and recall of associative memory in crickets. *Front. Behav. Neurosci.* **4**, 172 (2010).
36. Brauning, P. Suboesophageal DUM neurons innervate the principal neuropiles of the locust brain. *Phil. Trans. R. Soc. Lond. B* **332**, 221–240 (1991).
37. Wendt, B. & Homberg, U. Immunocytochemistry of dopamine in the brain of the locust *Schistocerca gregaria*. *J. Comp. Neurol.* **321**, 387–403 (1992).
38. Laurent, G. & Naraghi, M. Odorant-induced oscillations in the mushroom bodies of the locust. *J. Neurosci.* **14**, 2993–3004 (1994).
39. Hammer, M. An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees. *Nature* **366**, 59–63 (1993).
40. Gütig, R. & Sompolinsky, H. The tempotron: a neuron that learns spike timing-based decisions. *Nature Neurosci.* **9**, 420–428 (2006).
41. Huerta, R. & Nowotny, T. Fast and robust learning by reinforcement signals: explorations in the insect brain. *Neural Comput.* **21**, 2123–2151 (2009).
42. Seol, G. H. *et al.* Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron* **55**, 919–929 (2007).
43. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
44. Izhikevich, E. M. Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**, 1569–1572 (2003).
45. Mazor, O. & Laurent, G. Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* **48**, 661–673 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by the Lawrence Hanson Chair at Caltech, the National Institutes on Deafness and other Communication Disorders, Caltech's Broad Fellows Program, the Office of Naval Research (grants N00014-07-1-0741 and N00014-10-1-0735) and the Max Planck Society. We are grateful to L.-P. Mok for help with locust dissections and to E. Schuman, A. Siapas, E. Lubenov, M. Papadopoulou and members of the Laurent lab for comments on the manuscript.

Author Contributions S.C. and G.L. designed the experiments and simulations, discussed the results and wrote the paper. S.C. carried out the experiments and simulations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.L. (gilles.laurent@brain.mpg.de) or S.C. (stijn@caltech.edu).

METHODS

Preparation and stimuli. All results were obtained *in vivo* from locusts (*Schistocerca americana*) housed in an established, crowded colony. Young adults of either sex were immobilized in a holder. Both antennae were secured in place with respect to the olfactory delivery system and remained intact for olfactory stimulation. The brain was exposed, desheathed and superfused with locust saline, as previously described³⁸.

Odour delivery. Odours were diluted 10% v/v in paraffin oil. They were delivered by injection of a controlled volume of odorized air within a constant stream of desiccated air. Total airflow was set to 0.85 l min^{-1} and the odour was further diluted by one-third in air. Teflon tubing was used at and downstream of the mixing point to prevent odour lingering and cross-contamination.

Intradendritic recordings. Sharp electrode recordings from the dendrites of bLNs were made with borosilicate glass micropipettes (d.c. resistance, 150–200 M Ω) filled with 1 M potassium acetate. The cell type from which the data are derived could be recognized by several characteristics: responses to odour, subthreshold baseline activity profile and response to electrical stimulation of Kenyon cells¹⁷. Recordings from bLNs were always made from dendrites in the β -lobe. To assess changes in bLN instantaneous firing rate, bLN spike times were either binned per simultaneously recorded LFP cycle (Fig. 2f) or convolved with a 50-ms Gaussian kernel to generate smoothed PSTHs (Fig. 4c, e; shading, s.e.). EPSP slope was measured as the maximum slope over the 10–90% range of maximum amplitude. The average EPSP waveforms shown at bottom in Fig. 3a–c were computed from the same data as the slope measurements shown at top in Fig. 3a–c. The electrical stimulation rate was 0.1 s^{-1} (with no intertrial gap). The results in this work were obtained from 55 animals; in two instances the same animal was used for two distinct bLNs with OCT injection.

LFP recordings. LFPs were recorded in the mushroom body Kenyon cell soma layer using Michigan probes (<http://www.neuronexustech.com>).

Electrical stimulation. Twisted-wire tetrodes obtained from FHC (no. CE4B75) were modified for monopolar stimulation, with the casing serving as the anode. The tips of the tetrodes were splayed such that the distance between the exposed tips was approximately equal to 60% of the diameter of the mushroom body calyx. The exposed end of the stimulating electrode was embedded among Kenyon cell somata. The tetrodes were electroplated with gold solution to impedances of 200–350 k Ω at 1 kHz. Stimulating currents (5–140 μA , 0.1 ms) were generated using an STG1000 Multichannel System (as in ref. 17).

OCT injection. Patch pipettes were back-filled with a 100 μM OCT hydrochloride, 0.01% Fast Green solution. A pneumatic picopump (WPI) was used to apply one 50-ms, 5-p.s.i. pressure pulse per trial. Each pulse injected $\sim 5 \text{ pl}$ (as measured by previous injection into a drop of oil). The injected solution was localized to the β -lobe, as verified by bounded diffusion of the Fast Green dye.

Simulations. All simulations were carried out with networks of Izhikevich units⁴⁴. Networks consisted of 30 units each with dynamics given by

$$\frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I$$

$$\frac{du}{dt} = a(bv - u)$$

where $a = 0.02$, $b = 0.2$, $c = -65$, $d = 8$ and if $v \geq 30$ then $v \leftarrow c$ and $u \leftarrow u + d$. Synaptic interactions and external inputs were introduced through the I term, which for the j th unit takes the form

$$I_j = \sum_i w_{\text{bLN}}(i, j) f_{\text{bLN}}(i) + \sum_k w_{\text{KC}}(k, j) f_{\text{KC}}(k)$$

Here w_{bLN} is a matrix of inhibitory weights among network units, f_{bLN} is the vector of spike times generated by the network, w_{KC} is a matrix of excitatory input weights and f_{KC} is the vector of spike times that constitutes the input to the network. At every trial, f_{KC} is drawn from a distribution characterized by three experimental observations: odour-evoked Kenyon cell spiking phase distribution relative to the LFP¹², average odour-evoked Kenyon cell population PSTH⁴⁵ and Kenyon cell baseline activity¹². The matrix w_{bLN} is estimated from experimental recordings and kept fixed during the simulation. The matrix w_{KC} is drawn from a normal distribution at the onset of the simulation, with variable means for different initial conditions; it is updated in every trial according to

$$\Delta w = \begin{cases} \lambda f_{-}(w) K(\Delta t) & \text{if } \Delta t \leq 0 \\ \lambda f_{+}(w) K(\Delta t) & \text{if } \Delta t > 0 \end{cases}$$

The temporal filter, $K(\Delta t)$, is the experimentally derived STDP curve: either the unbiased Hebbian form as in ref. 17 or the depression-biased form that results from pairing with OCT as in Fig. 3f. The learning rate was set by the number of pairings carried out to arrive at the STDP curve in experiments. The weight dependences of the updating functions $f_{+}(w)$ and $f_{-}(w)$ are as follows:

$$f_{+}(w) = (1 - w)^{\mu}$$

$$f_{-}(w) = w^{\mu}$$

As described in ref. 24, the value of μ determines the shape of the equilibrium weight distribution. A value ($\mu = 0.08$) was chosen such that the weight distribution in the simulation at equilibrium matched the experimental EPSP size distribution (Supplementary Information, section 3). As described above, parameters in the model were constrained by experimental estimates, except for a , b , c and d , which characterize the dynamics of individual model units. These parameters took the typical values used for regular spiking neurons⁴⁴.

DNA breaks and chromosome pulverization from errors in mitosis

Karen Crasta^{1,2,3}, Neil J. Ganem^{1,2,3}, Regina Dagher^{1,2,3}, Alexandra B. Lantermann¹, Elena V. Ivanova⁴, Yunfeng Pan⁵, Luigi Nezi¹, Alexei Protopopov⁴, Dipanjan Chowdhury⁵ & David Pellman^{1,2,3}

The involvement of whole-chromosome aneuploidy in tumorigenesis is the subject of debate, in large part because of the lack of insight into underlying mechanisms. Here we identify a mechanism by which errors in mitotic chromosome segregation generate DNA breaks via the formation of structures called micronuclei. Whole-chromosome-containing micronuclei form when mitotic errors produce lagging chromosomes. We tracked the fate of newly generated micronuclei and found that they undergo defective and asynchronous DNA replication, resulting in DNA damage and often extensive fragmentation of the chromosome in the micronucleus. Micronuclei can persist in cells over several generations but the chromosome in the micronucleus can also be distributed to daughter nuclei. Thus, chromosome segregation errors potentially lead to mutations and chromosome rearrangements that can integrate into the genome. Pulverization of chromosomes in micronuclei may also be one explanation for ‘chromothripsis’ in cancer and developmental disorders, where isolated chromosomes or chromosome arms undergo massive local DNA breakage and rearrangement.

Whole-chromosome aneuploidy is a major feature of cancer genomes, yet its role in tumour development remains controversial^{1,2}. This contrasts with chromosome breaks and rearrangements, which are known to produce cancer-causing mutations. Recent genetic evidence demonstrates that increased rates of whole-chromosome mis-segregation can accelerate oncogenesis^{3–6}; however, the only established mechanism by which whole-chromosome segregation errors promote tumorigenesis is by facilitating the loss of heterozygosity for tumour suppressors⁷. Intriguingly, two animal models where whole-chromosome segregation errors result in robust tumour development also display extensive structural alterations in chromosomes^{6,8}. This raises the question of whether errors in mitosis can predispose to DNA damage. We considered the possibility that segregation of chromosomes into micronuclei might produce DNA damage. Whole-chromosome-containing micronuclei form from anaphase lagging chromosomes^{9–13}; micronuclei can also be generated from acentric fragments of chromosomes¹¹. Micronuclei have many features of primary nuclei, but much controversy surrounds their actual composition and functional properties. Studies differ on whether micronuclei are transcriptionally active, replicate DNA, mount a normal DNA damage response, or assemble normal nuclear envelopes; moreover, the ultimate fate of chromosomes trapped within micronuclei remains unclear^{11,14,15}.

DNA damage in micronuclei

To determine if newly formed whole-chromosome micronuclei develop DNA damage, we generated micronuclei in synchronized cells and tracked them through the cell cycle. As a first synchronization approach, micronuclei were generated in non-transformed RPE-1 and transformed U2OS cells by release from nocodazole-induced microtubule depolymerization. When mitotic cells are released from nocodazole, spindles reassemble abnormally, producing merotelic kinetochore attachments (one kinetochore attached to two opposite spindle poles), lagging chromosomes, and ~10% of cells with micronuclei¹⁶. Because

prolonged mitotic arrest causes DNA damage^{17–19}, RPE-1 cells were arrested for a short (6 h) interval with nocodazole. After release from the 6 h nocodazole block, neither the primary nuclei nor the newly formed kinetochore-positive micronuclei showed significant DNA damage during the subsequent G1 phase as measured by damage-dependent phosphorylation of the histone variant H2AX²⁰ (γ -H2AX foci formation; Fig. 1a–c and Supplementary Fig. 2a–f), TdT-mediated dUTP nick end labelling (TUNEL) to detect DNA breaks²¹ (Fig. 1d and Supplementary Fig. 2e, f), and visualization of DNA breaks with the comet assay²² (Supplementary Fig. 3d, e).

Because aneuploidy can produce a p53-dependent G1 cell-cycle arrest^{23,24}, p53 was silenced by RNA interference (RNAi) to allow us to monitor the fate of micronucleated RPE-1 cells at later stages of the cell cycle. As expected, after S-phase entry, low-level DNA damage was detected in both the micronuclei and the primary nucleus²⁵; however, in G2 phase cells, most micronuclei showed DNA damage whereas almost none was detected in the primary nucleus (Fig. 1a–d). Similar results were observed in U2OS cells (Supplementary Fig. 3) and in cells where merotelic kinetochore-microtubule attachments were generated by knockdown of the kinetochore-associated microtubule depolymerase MCAK²⁶ or the kinetochore protein NUF2 (ref. 27) (Supplementary Fig. 4a–c). MCAK knockdown does not delay cells in mitosis, demonstrating that the acquisition of DNA damage in micronuclei is independent of mitotic arrest^{17–19}. This damage did not represent activation of apoptosis because it was not accompanied by caspase-3 activation, and it was not suppressed by a pan-caspase inhibitor (Supplementary Fig. 5).

As a completely independent method for generating micronuclei, we used a human cell line (HT1080) carrying a human artificial chromosome (HAC) with a kinetochore that could be conditionally inactivated²⁸. In this system, kinetochore assembly on the HAC is blocked by washout of doxycycline from the medium; consequently, the HAC is unable to attach to the mitotic spindle and is left behind at anaphase,

¹Department of Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02115, USA. ²Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA. ³Howard Hughes Medical Institute, 450 Brookline Avenue, Boston, Massachusetts 02115, USA. ⁴Belfer Institute for Applied Cancer Science, Dana-Farber Cancer Institute, 4 Blackfan Street, Boston, Massachusetts 02115, USA. ⁵Department of Radiation Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02115, USA.

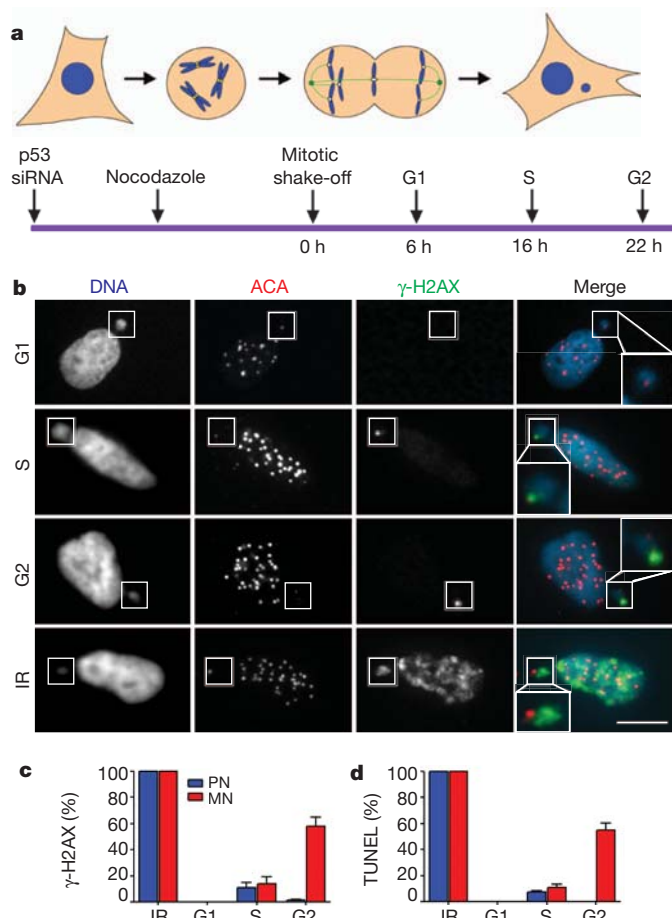


Figure 1 | Micronuclei from lagging chromosomes develop DNA breaks. **a**, Schematic of the experiment. **b**, Representative images of micronucleated G1 (6 h), S (16 h), G2 (22 h) and irradiated RPE-1 cells (IR) labelled for DNA (blue), centromeres (ACA, red) and γ -H2AX (green). Insets show enlarged images of micronuclei. Scale bars, 10 μ m. **c**, **d**, Percentage of primary nuclei (PN, blue bars) and micronuclei (MN, red bars) with γ -H2AX foci (**c**) and TUNEL labelling (**d**) (3 experiments, $n = 100$). Errors bars indicate s.e.m.

ultimately reforming as a micronucleus (Fig. 2a). Results obtained with the HAC-containing micronuclei (Fig. 2b–d) were in agreement with the results obtained by the other synchronization methods described above. Taken together, micronuclei do not exhibit significant DNA damage during G1, but a large fraction of micronuclei acquire DNA damage during S phase that persists into G2.

Defective DNA replication in micronuclei

To determine directly whether acquisition of DNA damage in micronuclei requires DNA replication, synchronized micronucleated cells were released into medium containing thymidine to block DNA replication. Blocking DNA replication with thymidine abolished the acquisition of DNA damage (Fig. 3a–c), demonstrating that the breaks in micronuclei occur in a replication-dependent manner.

We next tested whether the acquisition of DNA damage in micronuclei is mediated by aberrant DNA replication. This possibility was suggested because micronuclei in G2 cells that had γ -H2AX foci were almost always additionally labelled for phosphorylated RPA2 (Ser 33)²⁹, a marker of DNA replication stress (Supplementary Fig. 4d). Moreover, the characteristic 5-bromodeoxyuridine (BrdU) labelling patterns of early, mid and late S-phase cells³⁰ suggested that DNA replication in micronuclei is inefficient and asynchronous with the primary nucleus (Fig. 3d). By measuring the fluorescence intensity of BrdU relative to DNA area from single focal plane confocal images, DNA replication in micronuclei was detected at a markedly reduced level throughout S

phase when compared to the primary nuclei (Fig. 3e). Notably, pulse-labelling with BrdU demonstrated asynchronous DNA replication of the micronuclei relative to the primary nucleus, with ~30% of micronuclei replicating DNA in G2 cells (22 h; Fig. 3d).

The compromised DNA replication observed in newly formed micronuclei led us to test whether complexes necessary to license DNA replication origins are recruited to chromosomes in micronuclei. The first step in assembling replication origins is the loading of the origin recognition complex (ORC) after chromosome decondensation during telophase³¹. Quantitative fluorescence imaging suggested that ORC recruitment is equally efficient in the micronuclei as in the primary nucleus (Supplementary Fig. 6a). By contrast, newly generated micronuclei showed significant reduction in the recruitment of replicative DNA helicase components MCM2 and MCM3 (refs 32, 33) as well as the replication initiation factor CDT1 (ref. 31) (Fig. 4a, b and Supplementary Fig. 6b, c). We observed an increased recruitment of MCM subunits into micronuclei as cells progressed through the cell cycle, but the levels of MCM subunits never exceeded 20% of that observed in the primary nucleus (Fig. 4b).

To examine whether the DNA damage response in micronuclei was also abnormal, cells were treated with replication inhibitor (aphidicolin) or exposed to 2 Gy irradiation to induce damage and then labelled to detect components of the DNA damage response. After aphidicolin treatment, micronuclei showed clear signal for γ -H2AX and its kinase ATR (ataxia telangiectasia mutated-related), but downstream constituents of the DNA damage response such as 53BP1 were not efficiently recruited, as previously reported³⁴ (Supplementary Fig. 7). A similar defect in the recruitment of DNA-damage-response factors into micronuclei was also observed after irradiation (Supplementary Fig. 8a). This recruitment defect had clear consequences for DNA damage repair: whereas γ -H2AX foci were resolved in primary nuclei by 6 h after 2 Gy irradiation, γ -H2AX foci persisted in micronuclei for >24 h (Supplementary Fig. 8b). The failure of micronuclei to recruit MCM subunits and components of the DNA repair machinery may be due, at least in part, to a defect in nucleocytoplasmic transport. Newly formed micronuclei have a marked reduction in the density of nuclear pores^{11,14,15} (Supplementary Fig. 9) and exhibit strongly reduced nuclear import (Fig. 4c, d and Supplementary Fig. 10).

Micronuclei and chromosome pulverization

Next we tested the prediction that abnormal DNA replication in micronuclei could generate chromosome breaks. Chromosome spreads were prepared from non-transformed cells in the first cell cycle after release from nocodazole or from DMSO-treated controls (Fig. 5a–f and Supplementary Fig. 11a). Notably, 7.6% of chromosome spreads prepared from micronucleated cells released from nocodazole block ($n = 722$) exhibited chromosomes with a ‘pulverized’ appearance, characterized by fragments of 4',6-diamidino-2-phenylindole (DAPI)-stained material that were often clustered into a discrete location on the spread³⁵ (Fig. 5a, d (yellow arrows), i and Supplementary Fig. 11b). By contrast, pulverization was exceedingly rare (0.14%) in controls. Multicolour fluorescence chromosome painting (SKY) confirmed that 34 of 47 spreads with pulverized chromosomes were composed of fragments from a single chromosome and another eight were from two chromosomes (Fig. 5a–i and Supplementary Fig. 11). Despite this marked disruption to chromosomal structure, only one cell harboured non-reciprocal chromosome translocations 24 h after nocodazole release ($n = 101$), whereas no translocations were observed in control cells (data not shown). At 72 h after nocodazole release ($n = 100$), we still did not detect chromosome translocations, but we did detect truncated or derivative single chromosomes in three cells. Although our experiments show that whole-chromosome mis-segregation does not produce frequent non-reciprocal translocations, we do not exclude the possibility that the DNA damage acquired in micronuclei could result in translocations, albeit at a frequency lower than ~1/100 per generation.

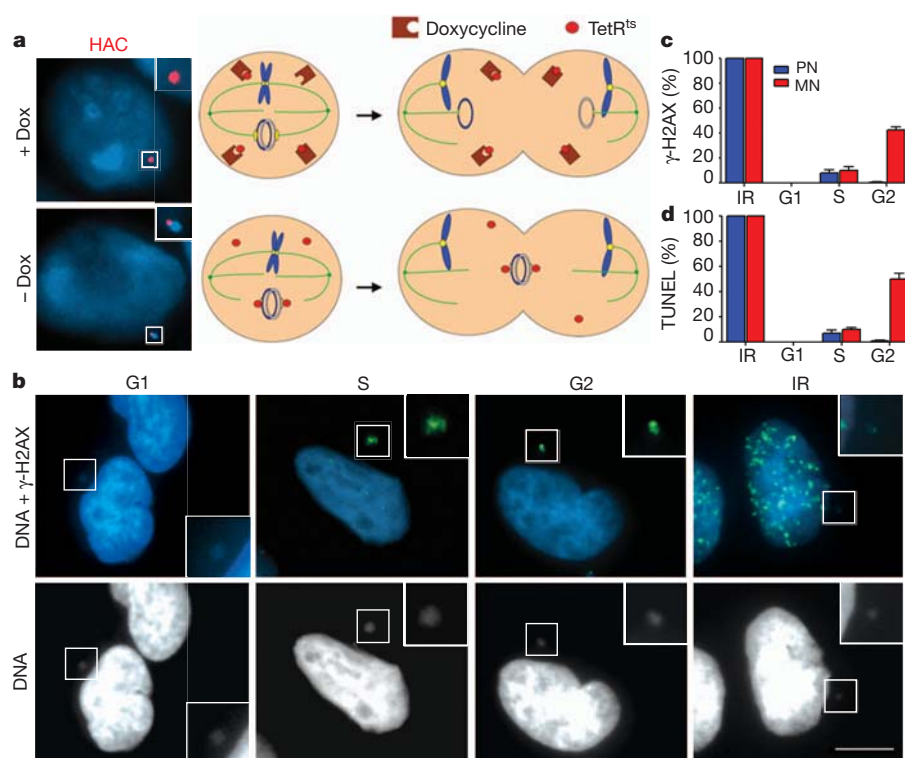


Figure 2 | DNA breaks in a HAC targeted to a micronucleus. **a**, Right: schematic; left: fluorescence *in situ* hybridization images of HAC (red) in a primary nucleus (+Dox) or micronucleus (-Dox). **b**, Images of micronucleated cells as in Fig. 1b (enlarged and brightened in insets). Scale bars,

10 μ m. **c**, **d**, Percentage of primary nuclei (blue bars) and centric micronuclei (red bars) with γ -H2AX foci (**c**) and TUNEL labelling (**d**) (3 experiments, $n = 100$). Errors bars indicate s.e.m.

Chromosome pulverization has been observed in cell-fusion experiments where chromosomes from an S-phase cell are pulverized because of exposure to signals from mitotic cytoplasm³⁶. It has also been observed when aberrant late-replicating chromosome translocations are generated³⁷. The mechanism of pulverization involves compaction of partially replicated chromosomes induced by mitotic cyclin-dependent kinase activity, termed premature chromosome compaction^{35,38,39}. By labelling micronucleated and control cells for 2 h with BrdU immediately before chromosome condensation and mitotic arrest, we found that 25 of 30 pulverized chromosomes were undergoing DNA replication (Fig. 5h). The fact that only micronuclei

incorporate BrdU in G2 cells (Figs 3d and 5g) demonstrates that the pulverized chromosomes are derived from micronuclei.

The fate of chromosomes in micronuclei

We next addressed the potential for chromosome aberrations acquired in micronuclei to be reincorporated into the genome. We used long-term live-cell imaging to determine the fate of spontaneously arising or experimentally induced micronuclei in a variety of cell lines (Supplementary Table 1). Imaging analysis revealed that the vast majority of micronuclei (525 of 541; 97%) are stably maintained during interphase. Although micronuclei may be extruded in some cells¹¹, extrusion was not detected in our imaging experiments. Furthermore, micronuclei were not visibly degraded¹¹, did not co-localize with lysosomes (Supplementary Fig. 12), and did not appear to fuse with the primary nucleus. By contrast, our experiments indicated that after nuclear envelope breakdown some micronuclei might join the other mitotic chromosomes and be distributed to daughter nuclei. Of the 331 micronucleated cells that were tracked as they underwent cell division, 16% (53 of 331) gave rise to daughter cells

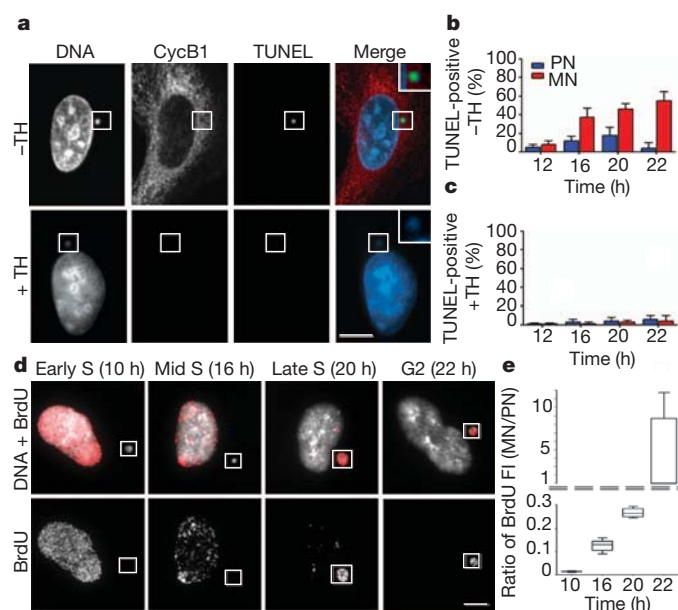


Figure 3 | DNA damage in micronuclei results from aberrant DNA replication. **a–c**, DNA replication requirement for acquisition of DNA damage in micronuclei. **a**, RPE-1 cells were synchronized as in Fig. 1a and released into media with (+TH) or without (-TH) 2 mM thymidine. Cells were co-labelled for TUNEL (green) and cyclin B1 (red). Shown are images from the 22 h sample. Scale bar, 10 μ m. **b**, **c**, Percentage of TUNEL-positive primary nuclei (blue bars) and micronuclei (red bars) with (c) or without (b) thymidine treatment; (3 experiments, $n = 100$). Errors bars indicate s.e.m. **d**, Inefficient and asynchronous DNA replication in micronuclei. RPE-1 cells as in Fig. 1a were pulse-labelled with BrdU and labelled: DNA (white) and BrdU (red). **e**, The ratio of BrdU incorporation in micronuclei relative to primary nuclei after a 30-min pulse label. Normalized fluorescence intensity (FI) measurements are as shown in the box and whisker plots (2 experiments, $n = 50$). Boxes represent upper and lower quartiles, lines within boxes represent median, and the error bars comprise the whiskers which extend to the maximum and minimum value data sets.

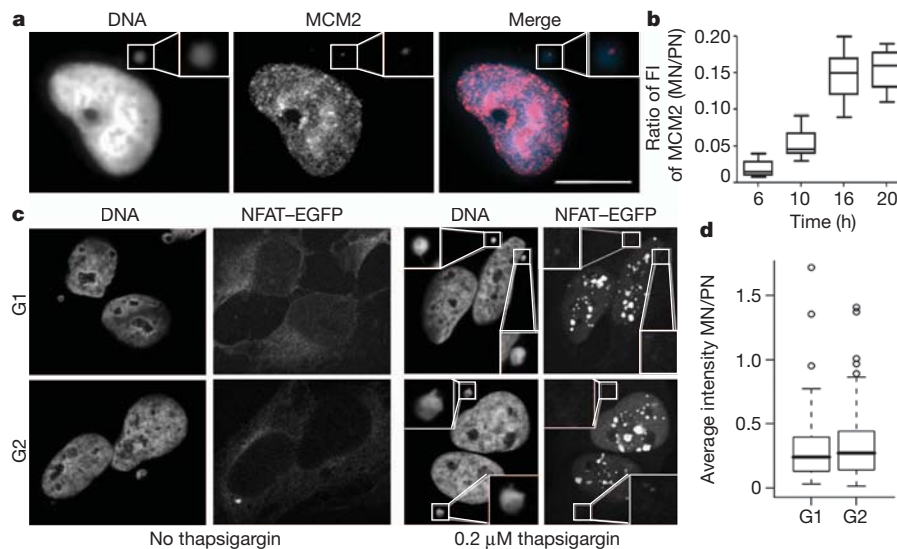


Figure 4 | Defective MCM2-7 complex recruitment, DNA damage response and nucleocytoplasmic transport in micronuclei. **a**, Impaired MCM2 recruitment into micronuclei. G1 RPE-1 cells were synchronized as in Fig. 1a and stained for chromatin-bound MCM2 and DNA. Scale bar, 10 μ m. **b**, Relative MCM2 fluorescence intensity (FI) as in Fig. 3e. Approximate cell cycle stage of time points: 6 h, G1; 10 h, early S phase; 16 h, mid S phase; 20 h, late S phase; (2 experiments, $n = 50$). **c**, **d**, Micronuclei are partially defective for

nuclear import of NFATc1-EGFP. **c**, Representative images of micronucleated U2OS cells stably expressing NFATc1-EGFP in G1 and G2 with or without treatment with 0.2 μ M thapsigargin for 10 min. **d**, Ratio of NFATc1-EGFP fluorescence intensity between micronuclei and primary nuclei; (3 experiments, $n = 30$). For **b** and **d**, error bars comprise the whiskers which extend to the maximum and minimum value data sets.

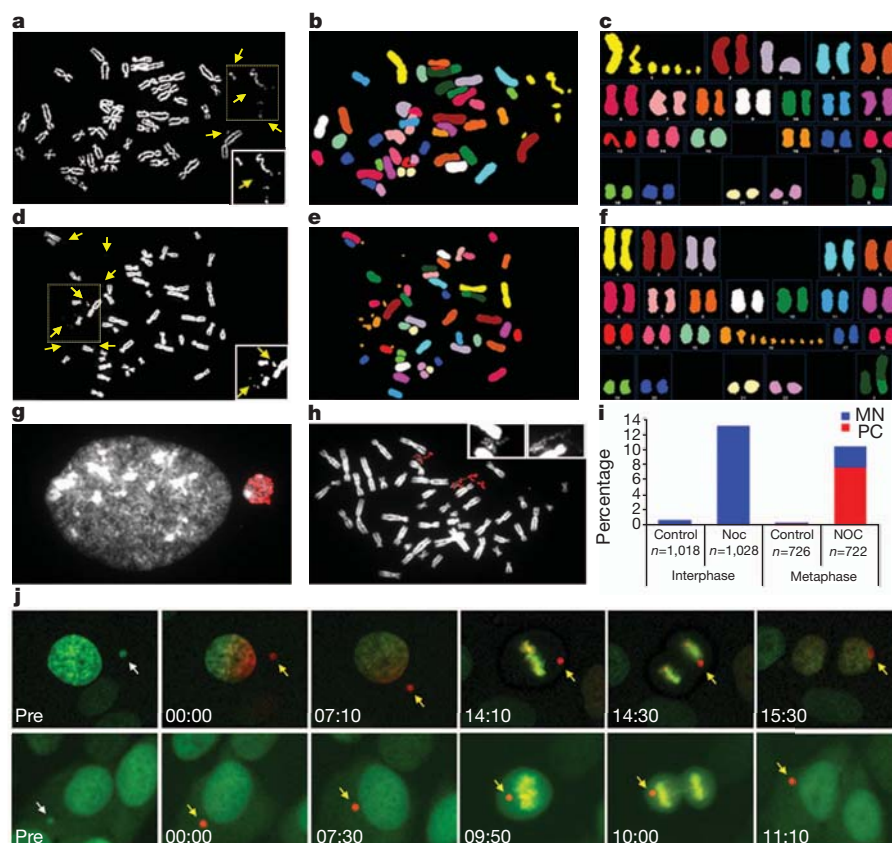


Figure 5 | The fate of chromosomes in micronuclei. Micronuclei were induced in RPE-1 cells as described in Fig. 1a (after p53 knockdown) and chromosome spreads were prepared 24 h after nocodazole release. **a–c**, Pulverization of chromosome 1 demonstrated by DAPI staining (chromosome fragments, brightened in inset, yellow arrows) (**a**), SKY probes (pseudo-coloured) (**b**), and aligned SKY karyotype (**c**). **d–f**, Pulverization of chromosome 16, viewed as in **a–c**. **g**, A BrdU-positive micronucleus-containing G2 RPE-1 cell. 2 h BrdU pulse label. DNA, white; BrdU, red. **h**, Selective BrdU labelling of a pulverized chromosome. **i**, The percentage of cells with intact

micronuclei (blue bars) or pulverized chromosomes (PC, red bars) from control or nocodazole-released (NOC) RPE-1 cells. Interphase is 18-h sample and metaphase is 24 h. **j**, Images from live-cell experiment showing a pre-converted green fluorescent micronucleus (white arrows) photo-converted to a red-fluorescent micronucleus (yellow arrows) imaged through mitosis. After anaphase, the micronucleus either reincorporated with the primary nucleus or failed to reincorporate and reformed as a micronucleus. Top row: reincorporation of micronuclei into primary nucleus; bottom row: no reincorporation. Time is shown as h:min.

that lacked micronuclei (Supplementary Table 1 and Supplementary Movies 1 and 2).

To test directly whether micronuclei reincorporate during mitosis, we generated a U2OS cell line stably expressing H2B fused to the photoconvertible fluorescent protein Kaede⁴⁰, which converts from green to red fluorescence emission after ultraviolet illumination. This enabled us to photoactivate selectively and subsequently track a single chromosome encapsulated in a micronucleus throughout mitosis (Fig. 5j and Supplementary Movies 3 and 4). This experiment demonstrated that chromosomes within micronuclei reincorporate into daughter nuclei at a significant frequency (11 of 29; 38%). The remaining micronuclei persisted in cells well into the second generation.

Discussion

Together, our findings indicate a novel mechanism by which whole-chromosome segregation errors can cause chromosome breaks and potentially cancer-causing mutations. Mitotic errors can lead to lagging chromosomes which in turn can be partitioned into micronuclei¹³. This can produce DNA damage in two ways. First, DNA damage can result as a direct consequence of aberrant DNA replication, potentially due to a reduced density of replication origins. Second, chromosome pulverization due to premature chromosome compaction can occur if a micronucleated cell enters mitosis with micronuclei still undergoing DNA replication. This may not be the only mechanism by which whole-chromosome mis-segregation can cause DNA damage, as it was recently proposed that lagging chromosomes can be damaged by the cytokinetic furrow⁴¹. In contrast to previous studies that have suggested that most micronuclei are inactive, discarded or degraded¹¹, we find that micronuclei persist over several generations but that the chromosomes within the micronuclei can be segregated at a significant frequency into daughter cell nuclei. Thus, DNA rearrangements and mutations acquired in micronuclei could be incorporated into the genome of a developing cancer cell.

Our experiments suggest a mechanism that could explain the recently discovered phenomenon of chromothripsis^{42,43}. Chromothripsis was discovered by cancer genome sequencing and DNA copy number analysis⁴². It is defined by small-scale DNA copy number changes and extensive intrachromosomal rearrangements that are restricted to a single chromosome or chromosome arm. Two non-exclusive models have been proposed for chromothripsis are: (1) fragmentation of a chromosome followed by stitching together of the resulting fragments by non-homologous end joining⁴²; and (2) aberrant DNA replication resulting in fork stalling and template switching or microhomology-mediated break-induced replication^{43,44}. Partitioning of a chromosome into a spatially isolated micronucleus can explain why a single chromosome would be subject to aberrant DNA replication and can explain how extensive DNA damage can be restricted to a single chromosome. Because micronuclei may persist as discrete entities over several generations, damage to the chromosome need not occur all at once^{42,43}, but could accrue from multiple failed attempts at DNA replication. Chromothripsis restricted to a chromosome arm could be generated from acentric chromosome fragments, as chromosome pulverization was previously correlated with the presence of micronuclei in cancer cell lines after DNA damage^{35,45}. Although chromothripsis seems to be relatively rare, DNA damage in micronuclei could be more common. We propose that chromothripsis may be an extreme outcome of a mutagenesis mechanism that could be widespread in human cancer.

METHODS SUMMARY

All cell lines were maintained at 37 °C with 5% CO₂ atmosphere. To generate micronuclei, cells were treated with nocodazole (100 ng ml⁻¹) for 6 h followed by mitotic shake-off and nocodazole washout. Alternatively, micronuclei were generated by depletion of MCAK or NUF2 by short interfering RNA (siRNA). Micronuclei were also generated in HT1080 cells by inactivation of the kinetochore of the human artificial chromosome by washout of doxycycline. Detailed descriptions of immunofluorescence, RNAi, BrdU labelling, comet assays, inactivation of the human artificial chromosomes (HAC), flow cytometry, quantitative analysis of

nuclear import, chromosome spreads, spectral karyotyping, antibodies used, photoconversion of H2B-Kaede and long-term live-cell imaging experiments can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 August; accepted 20 December 2011.

Published online 18 January 2012.

- Gordon, D. J., Resio, B. & Pellman, D. Causes and consequences of aneuploidy in cancer. *Nature Rev. Genet.* <http://dx.doi.org/10.1038/nrg3123> (in the press).
- Weaver, B. A. & Cleveland, D. W. Aneuploidy: instigator and inhibitor of tumorigenesis. *Cancer Res.* **67**, 10103–10105 (2007).
- Schvartzman, J. M., Sotillo, R. & Benezra, R. Mitotic chromosomal instability and cancer: mouse modelling of the human disease. *Nature Rev. Cancer* **10**, 102–115 (2010).
- Ricke, R. M., van Ree, J. H. & van Deursen, J. M. Whole chromosome instability and cancer: a complex relationship. *Trends Genet.* **24**, 457–466 (2008).
- Weaver, B. A., Silk, A. D., Montagna, C., Verdier-Pinard, P. & Cleveland, D. W. Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* **11**, 25–36 (2007).
- Sotillo, R. et al. Mad2 overexpression promotes aneuploidy and tumorigenesis in mice. *Cancer Cell* **11**, 9–23 (2007).
- Baker, D. J., Jin, F., Jeganathan, K. B. & van Deursen, J. M. Whole chromosome instability caused by Bub1 insufficiency drives tumorigenesis through tumor suppressor gene loss of heterozygosity. *Cancer Cell* **16**, 475–486 (2009).
- Fujiwara, T. et al. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
- Cimini, D. Merotelic kinetochore orientation, aneuploidy, and cancer. *Biochim. Biophys. Acta* **1786**, 32–40 (2008).
- Ganem, N. J., Godinho, S. A. & Pellman, D. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460**, 278–282 (2009).
- Terradas, M., Martin, M., Tusell, L. & Genesca, A. Genetic activities in micronuclei: is the DNA entrapped in micronuclei lost for the cell? *Mutat. Res.* **705**, 60–67 (2010).
- Gisselsson, D. et al. Abnormal nuclear shape in solid tumors reflects mitotic instability. *Am. J. Pathol.* **158**, 199–206 (2001).
- Thompson, S. L. & Compton, D. A. Chromosome missegregation in human cells arises through specific types of kinetochore-microtubule attachment errors. *Proc. Natl Acad. Sci. USA* **108**, 17974–17978 (2011).
- Hoffelder, D. R. et al. Resolution of anaphase bridges in cancer cells. *Chromosoma* **112**, 389–397 (2004).
- Géraud, G. et al. Three-dimensional organization of micronuclei induced by colchicine in PtK1 cells. *Exp. Cell Res.* **181**, 27–39 (1989).
- Cimini, D. et al. Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J. Cell Biol.* **153**, 517–528 (2001).
- Dalton, W. B., Nandan, M. O., Moore, R. T. & Yang, V. W. Human cancer cells commonly acquire DNA damage during mitotic arrest. *Cancer Res.* **67**, 11487–11492 (2007).
- Quignon, F. et al. Sustained mitotic block elicits DNA breaks: one-step alteration of ploidy and chromosome integrity in mammalian cells. *Oncogene* **26**, 165–172 (2007).
- Orth, J. D., Loewer, A., Lahav, G. & Mitchison, T. J. Prolonged mitotic arrest induces partial activation of apoptosis, DNA damage and p53 activation. *Mol. Biol. Cell* <http://dx.doi.org/10.1091/mbc.E11-09-0781> (2011).
- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
- Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. Identification of programmed cell death *in situ* via specific labeling of nuclear DNA fragmentation. *J. Cell Biol.* **119**, 493–501 (1992).
- Singh, N. P., McCoy, M. T., Tice, R. R. & Schneider, E. L. A simple technique for quantitation of low levels of DNA damage in individual cells. *Exp. Cell Res.* **175**, 184–191 (1988).
- Thompson, S. L. & Compton, D. A. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J. Cell Biol.* **188**, 369–381 (2010).
- Li, M. et al. The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. *Proc. Natl Acad. Sci. USA* **107**, 14188–14193 (2010).
- Branzei, D. & Foiani, M. Maintaining genome stability at the replication fork. *Nature Rev. Mol. Cell Biol.* **11**, 208–219 (2010).
- Kline-Smith, S. L., Khodjakov, A., Hergert, P. & Walczak, C. E. Depletion of centromeric MCAK leads to chromosome congression and segregation defects due to improper kinetochore attachments. *Mol. Biol. Cell* **15**, 1146–1159 (2004).
- DeLuca, J. G., Moree, B., Hickey, J. M., Kilmarin, J. V. & Salmon, E. D. hNuf2 inhibition blocks stable kinetochore-microtubule attachment and induces mitotic cell death in HeLa cells. *J. Cell Biol.* **159**, 549–555 (2002).
- Nakano, M. et al. Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* **14**, 507–522 (2008).
- Binz, S. K., Sheehan, A. M. & Wold, M. S. Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair* **3**, 1015–1024 (2004).
- O'Keefe, R. T., Henderson, S. C. & Spector, D. L. Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. *J. Cell Biol.* **116**, 1095–1110 (1992).
- DePamphilis, M. L. & Bell, S. D. *Genome Duplication* (Garland Science, 2011).

32. Dimitrova, D. S., Prokhorova, T. A., Blow, J. J., Todorov, I. T. & Gilbert, D. M. Mammalian nuclei become licensed for DNA replication during late telophase. *J. Cell Sci.* **115**, 51–59 (2002).
33. Mendez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* **20**, 8602–8612 (2000).
34. Terradas, M., Martin, M., Tusell, L. & Genesca, A. DNA lesions sequestered in micronuclei induce a local defective-damage response. *DNA Repair* **8**, 1225–1234 (2009).
35. Obe, G., Beek, B. & Vaidya, V. G. The human leukocyte test system. III. Premature chromosome condensation from chemically and x-ray induced micronuclei. *Mutat. Res.* **27**, 89–101 (1975).
36. Johnson, R. T. & Rao, P. N. Mammalian cell fusion: induction of premature chromosome condensation in interphase nuclei. *Nature* **226**, 717–722 (1970).
37. Smith, L., Plug, A. & Thayer, M. Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations. *Proc. Natl Acad. Sci. USA* **98**, 13300–13305 (2001).
38. Nichols, W. W., Levan, A., Aula, P. & Norrby, E. Chromosome damage associated with the measles virus *in vitro*. *Hereditas* **54**, 101–118 (1965).
39. Kato, H. & Sandberg, A. A. Chromosome pulverization in human binucleate cells following colcemid treatment. *J. Cell Biol.* **34**, 35–45 (1967).
40. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H. & Miyawaki, A. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl Acad. Sci. USA* **99**, 12651–12656 (2002).
41. Janssen, A., van der Burg, M., Suzhai, K., Kops, G. J. & Medema, R. H. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science* **333**, 1895–1898 (2011).
42. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
43. Liu, P. *et al.* Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* **146**, 889–903 (2011).
44. Hastings, P. J., Lupski, J. R., Rosenberg, S. M. & Ira, G. Mechanisms of change in gene copy number. *Nature Rev. Genet.* **10**, 551–564 (2009).
45. Sen, S., Hittelman, W. N., Teeter, L. D. & Kuo, M. T. Model for the formation of double minutes from prematurely condensed chromosomes of replicating micronuclei in drug-treated Chinese hamster ovary cells undergoing DNA amplification. *Cancer Res.* **49**, 6731–6737 (1989).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. D'Andrea, M. E. McLaughlin, T. A. Rapoport, J. Walters and Pellman laboratory members for discussions and/or comments on the manuscript; L. Cameron for advice and help on microscopy; H. Li for irradiation of cells; M. Nitta and M. Hennessy for preliminary experiments; and V. Larionov, B. Stillman, J. Ellenberg, I. Mattaj, A. Miyawaki and T. Kuroda for reagents. D.P. was supported by the Howard Hughes Medical Institute and the NIH (GM083299); K.C. was a fellow of A*STAR Singapore; N.J.G. was a fellow of the Leukemia and Lymphoma Society; Y.P. and D.C. were funded by the NIH (1R01CA142698-01).

Author Contributions D.P. conceived the project; K.C., N.J.G., R.D., A.B.L., D.C. and D.P. designed the experiments; D.P., K.C. and N.J.G. wrote the manuscript with edits from all authors; K.C. contributed Figs. 1–3, 4a, b and Supplementary Figs 2–8 with help from R.D. N.J.G. contributed Fig. 5j, Supplementary Fig. 12, Supplementary Table 1 and Supplementary Movies. R.D., E.V.I. and A.P. contributed Fig. 5a–i and Supplementary Figs 1 and 11; A.B.L. contributed Fig. 4c, d and Supplementary Figs 9 and 10; Y.P. and D.C. contributed Supplementary Fig. 3d; L.N. contributed Fig. 2a.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to david_pellman@dfci.harvard.edu.

METHODS

Cell culture. All cell lines were maintained at 37°C with 5% CO₂ atmosphere in DMEM (U2OS, HT1080), DMEM:F12 (hTERT RPE-1), or DMEM:F12 medium without phenol red (H2B-GFP and H2B-mRFP U2OS, H2B-GFP and H2B-mRFP RPE-1, H2B-GFP Caco2, H2B-GFP SCC-114, H2B-GFP HeLa, H2B-mRFP MDA-231). All media were supplemented with 10% FBS, 100 IU ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin. HT1080 cells bearing the aliphoid^{tetO} human artificial chromosome and expressing the tetracycline repressor TetR fused to the transcriptional silencer tTS (a gift from V. Larionov) were also supplemented with 1 µg ml⁻¹ doxycycline (Sigma), 0.5 mg ml⁻¹ G418 (geneticin, Invitrogen) and 4 µg ml⁻¹ blasticidin S (Invitrogen). U2OS cells stably expressing human NFATc1 fused to the C terminus of EGFP (Thermo Scientific; R04-017-01) were supplemented with 0.5 mg ml⁻¹ G418.

Generation of whole-chromosome micronuclei. To generate cells with whole-chromosome micronuclei, hTERT RPE-1 and U2OS cells were treated with 100 ng ml⁻¹ nocodazole (Sigma) for 6 h. Mitotic cells were collected, washed twice with fresh medium containing 10% FBS, and then plated into medium containing 20% FBS where they completed cell division. To overcome the aneuploidy-induced p53-dependent G1 arrest associated with nocodazole washout experiments, RPE-1 cells were transfected with 50 nmol p53 siRNA (Smartpool, Dharmacon) using Lipofectamine RNAi Max (Invitrogen) 12 h before the nocodazole treatment, enabling cell-cycle progression of the subsequent aneuploid daughter cells with micronuclei. For NUF2 and MCAK knockdown, U2OS cells were serum starved for 2 days before transfection with 50 nmol NUF2 or MCAK siRNAs (Smartpool, Dharmacon). Cells were then washed in fresh medium containing 10% FBS and 24 h later treated with 4 µM dihydrocytochalasin B (DCB; Sigma) for 16 h to inhibit cytokinesis. The subsequent arrested tetraploid cells depleted of NUF2 or MCAK (many of which harboured micronuclei) were then treated for 6 h with 10 µM SB203580 (Sigma), a p38 inhibitor that promotes cell-cycle progression of tetraploids. Sequence information of the small interference RNA (siRNA) pools used from Dharmacon are as follows: human TP53 ON-TARGETplus SMARTpool siRNA L-003329-00-0005, (J-003329-14) 5'-GAAUUUGCGUGUGAGUA-3', (J-003329-15) 5'-GUGCAGCUGUGGUUGAUU-3', (J-003329-16) 5'-GCAGUCAGAUCCUAGCGUC-3', (J-003329-17) 5'-GGAGAAUUAUCCUCCUUC-3'; human KIF2C/MCAK ON-TARGETplus SMARTpool siRNA L-004955-00-0005, (J-004955-06) 5'-GGCAUAAGCUCCUGUGAAU-3', (J-004955-07) 5'-CCAACGCAGUAAUGGUUA-3', (J-004955-08) 5'-GCAAGCAACAGGUGCAAGU-3', (J-004955-09) 5'-UGACUGAUCCUAUCGAAGA-3'; human CDCA1/NUF2 ON-TARGETplus SMARTpool siRNA L-005289-00-0005, J-005289-06 5'-GAACGAGUAAACCACAAUA-3', (J-005289-07) 5'-UAGCUGAGAUUGUGAUUA-3', (J-005289-08) 5'-GGAUUGCAAUAAGUUCAA-3', (J-005289-09) 5'-AAACGAUAGUGCUGCAAGA-3'.

Generation of micronuclei containing the HAC. HT1080 cells carrying the HAC were cultured in doxycycline-free medium and treated with 100 ng ml⁻¹ nocodazole for 6 h to synchronize cells in mitosis. The absence of doxycycline enables TetR binding to TetO, which induces inactivation of the HAC centromere by tTS and the subsequent formation of a micronucleus after release from mitotic arrest. Cells released from mitotic arrest were synchronized at G0/G1 by serum starvation for 24 h, then released into the cell cycle with fresh medium containing 10% serum for further analysis.

Flow cytometry. Cells were fixed with 70% ethanol at 4°C followed by incubation with 250 µg ml⁻¹ RNase A and 10 µg ml⁻¹ propidium iodide at 37°C for 30 min. FACS analysis was performed with a FACScalibur flow cytometer (Becton Dickinson) and data analysed with CellQuest software.

Indirect immunofluorescence microscopy. For most experiments, cells were seeded on glass coverslips, washed in CSK buffer (10 mM PIPES, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂), pre-extracted with CSK buffer containing 0.5% Triton X-100 for 5 min at 4°C, and fixed in PBS containing 4% paraformaldehyde for 15 min. After fixation, cells were permeabilized in PBS-0.2% Triton X-100 for 5 min, blocked in blocking buffer (PBS containing 5% FBS, 2% BSA and 0.1% Triton X-100) for 30 min, and then incubated with primary antibodies at room temperature for 1 h or overnight at 4°C. Cells were washed with PBS-0.1% Triton X-100 and incubated with fluorescence-conjugated secondary antibodies (1:1,000, Molecular Probes) at room temperature for 1 h. Cells were also stained for DNA with Hoechst 33342 (1:5,000, Invitrogen) in PBS. More specific immunostaining protocols were as follows: for γ-H2AX and other DNA damage-response proteins, as in ref. 46; for ELYS and mAb414, ref. 47; for cyclin B1/γ-H2AX/BrdU, ref. 48, and for MCM2, CDT1 and ORC2, ref. 49. Images for most experiments were collected with a Yokogawa CSU-22 spinning disk confocal mounted on a Zeiss Axiovert microscope using 488, 561 and 640 nm laser light. A series of 0.3 µm optical sections was acquired using a ×100 1.4 NA Plan Apo objective with an Orca ER CCD camera (Hamamatsu Photonics). For nuclear pore quantification, images were collected on a Leica SP5 laser scanning

confocal microscope with a 405 nm laser and a white light laser tuned to 488 nm, 568 nm and 647 nm using a ×63 1.4 NA Plan Apo objective. Z-stacks were collected with a 0.3 µm step size with pinhole at 1 Airy unit. Photomultipliers were calibrated to ensure linear range. For quantification of MCM2 and BrdU signals, mean fluorescence intensity within the primary nuclei or micronuclei was measured using ImageJ software. For quantification of γ-H2AX, primary nuclei were classified as positive when five or more foci were detected, with the threshold of foci intensity set against the irradiated sample. The same threshold was applied to score γ-H2AX foci in micronuclei. Acquisition parameters, shutters, filter positions and focus were controlled by Slidebook software (Intelligent Imaging Innovations). Images presented in figures are maximum intensity projections of entire z-stacks, unless otherwise stated.

Antibodies for immunofluorescence. Samples were incubated with primary antibodies for human anti-centromere ACA (1:1,000; Antibodies, Inc.), rabbit γ-H2AX-Ser 139 (1:500, Cell Signaling), mouse γ-H2AX (clone JBW301, 1:500, Upstate), rabbit 53BP1 (1:500, Cell Signaling), rabbit phospho-53BP1-Ser 1778 (1:500, Cell Signaling), rabbit phospho-CHK1-Ser 317 (1:500, Cell Signaling), rabbit phospho-ATM-Ser 1981 (1:500, Cell Signaling), rabbit ATR (1:500, Cell Signaling), rabbit MRE11 (1:300, Abcam), rabbit RAD51 (1:100, Santa Cruz), mouse BRCA1 (1:300, Santa Cruz), mouse RPA2 (1:500; Abcam), rabbit phospho-RPA2-S33 (1:500, Bethyl Laboratories), MCM2 and ORC2 (1:300, gifts from B. Stillman), rabbit CDT1 (1:500, Bethyl Laboratories), rabbit geminin (1:500; Santa Cruz), mouse monoclonal antibody 414 (1:5,000, Abcam), ELYS/MEL28 (1:1,000, gift from I. Mattaj), mouse lamin A/C (1:300, Abcam), rabbit cyclin A (1:300; Santa Cruz), rabbit cyclin B1 (1:500; Santa Cruz), rabbit cyclin D1 (1:300; Santa Cruz) and mouse BrdU (Sigma). Secondary antibodies used were Alexa Fluor 488 (green), 594 (red) and 647 (far red) from Molecular Probes.

TUNEL assay. TUNEL assay was performed according to the manufacturer's instructions (*In situ* cell death detection kit, Roche).

BrdU labelling. Cells were pre-labelled with 10 µM BrdU for 30 min and subsequently fixed and permeabilized according to the manufacturers' conditions (5-bromo-2-deoxyuridine labelling and detection kit 1, Roche). Cells were then stained with anti-BrdU (1 h, diluted 1:300) and processed for immunofluorescence. In the box and whiskers plot, the box represents upper and lower quartiles, line within box represents median and the whiskers extend to the highest and lowest value data sets.

Long-term live-cell imaging. Labelled cells were grown on glass-bottom 12-well tissue culture dishes (Mattek) and imaged on either a Nikon TE2000-E2 or Nikon Ti-E inverted microscope. Both microscopes were equipped with a cooled CCD camera (TE2000, Orca ER, Hamamatsu; Ti-E, Coolsnap HQ2, Photometrics), a precision motorized stage (Bioprecision, Ludl), and Nikon Perfect Focus, and both were enclosed within temperature and CO₂-controlled environments that maintained an atmosphere of 37°C and 3–5% humidified CO₂. GFP or RFP images were captured at multiple points every 5–10 min for 3–5 days with either ×10 (0.3 NA) or ×20 (0.5, 0.75 NA) objectives, and exposure to fluorescent light was minimized to the greatest extent possible (all image acquisition used neutral density filters and/or 2 × 2 binning). Cell viability was confirmed by the continuous observation of mitotic cells throughout the duration of experiments. Captured images from each experiment were analysed using NIS-Elements software.

Imaging of H2B-Kaede. Tandem repeats of Kaede (a gift of A. Miyawaki) were fused to H2B and cloned into the pLenti6/V5 lentiviral vector (Invitrogen). A U2OS cell line stably expressing H2B-Kaede was generated by lentiviral infection followed by repeated FACS sorting of GFP⁺ cells. For tracking experiments, pre-converted micronuclei (GFP⁺) were identified and photoconverted with a brief (1–3 s) pulse of 350–400 nm ultraviolet light from a DAPI filter cube. Micronuclei that efficiently converted (RFP⁺) were subsequently imaged through the completion of the following mitosis using the same microscope setup as described above. Images were acquired every 10 min with a ×20 objective and 2 × 2 binning.

Chromosome spreads/spectral karyotyping. RPE-1 cells were treated with p53 siRNA (as previously described) for 16 h and then grown in medium supplemented with nocodazole (100 ng ml⁻¹) or vehicle control (DMSO) for 6 h. Mitotic cells were collected by shake-off, washed thoroughly in fresh medium without nocodazole, and re-plated in culture medium with 20% FBS to complete mitosis. FACS analysis demonstrated that the subsequent daughter cells progressed to S/G2-phase 18–20 h after release. At this point, interphase samples were collected and fixed to observe the fraction of cells containing micronuclei. In parallel, 100 ng ml⁻¹ colcemid (Invitrogen) was added to the DMSO or nocodazole-treated cells for an additional 6 h and then fixed. All samples were similarly processed for fixation: cells were pelleted and re-suspended in a hypotonic solution of 0.075 M KCl for 18 min, fixed in Carnoy's fixative (3:1 methanol:glacial acetic acid) and washed four times with Carnoy's fixative. All fixed samples were spread on slides for staining or hybridization. Mitotic samples were stained with Giemsa or Hoechst to visualize chromosomal abnormalities. Spectral karyotyping (SKY) was performed

on interphase and mitotic samples (from both DMSO- and nocodazole-treated samples) according to the SkyPaint DNA kit H-5 for human chromosomes procedure (Applied Spectral Imaging, SKY000029) and imaged on a Nikon Eclipse E6000 microscope equipped with the SD300 Spectracube and Spectral Imaging acquisition software. To determine if pulverized chromosomes were derived from late-replicating chromosomes in micronuclei, the above procedure was also performed with addition of BrdU labelling reagent (Roche Applied Science) only in the last 2 h of colcemid treatment before fixation. Mitotic samples from the BrdU samples were treated according to the 5-bromo-2'-deoxy-uridine labelling and detection kit I protocol (Roche Applied Science, 11296736001), except that the anti-BrdU antibody was diluted 1:50.

Single-cell gel electrophoresis (comet) assay. Single-cell comet assays were performed according to the manufacturer's instructions (Trevigen). Briefly, U2OS cells were collected at G1, S and G2 phases of the cell cycle after nocodazole washout was used to induce micronuclei. As controls, G1 cells were irradiated at 10 Gy. Cells were re-suspended in cold PBS at 2×10^5 cells ml⁻¹, mixed with low-melt agarose (1:10 ratio) and spread on frosted glass slides. After the agarose solidified, the slides were sequentially placed in lysis and alkaline solutions (Trevigen). Slides were then subjected to electrophoresis at 12 V for 10 min in 1× TBE buffer, fixed with 70% ethanol, and stained with DAPI. Nuclei were visualized using epifluorescent illumination on a Zeiss microscope and images were analysed with the NIH Image J program. DNA damage was quantified for 50 cells with micronuclei for each experimental condition by determining the tail DNA percentage using Comet Score (TriTek) software.

Quantitative analysis of nuclear pore complexes. To determine nuclear pore complex density in primary nuclei and micronuclei, cells were pre-extracted, fixed, permeabilized and immunostained with monoclonal 414 and ELYS antibodies. Confocal images were collected on a Leica SP5 laser scanning confocal with both a 405-nm and white light laser (at 488nm and 568nm) using a ×63 Plan Apo 1.4 NA oil objective. Z-stacks were collected with a 0.3 µm step size with pinhole at 1 Airy unit. Photomultipliers were calibrated to ensure linear range. Single layers of the Z-stack showing optimal nuclear pore staining for the primary nucleus and the micronuclei were chosen for quantitative analysis using MetaMorph. The areas of primary nuclei and micronuclei were judged from Hoechst staining and used to

determine the perinuclear rim area. Total fluorescence intensity of monoclonal antibody 414 and ELYS was measured in the perinuclear rim area. To compare nuclear pore densities between primary nuclei and micronuclei, the total nuclear pore fluorescence was divided by the perinuclear rim area of primary nuclei and micronuclei, resulting in an average intensity ratio.

Nuclear import. To measure nuclear import, we used U2OS cells stably expressing the fusion protein NFATc1-EGFP (Thermo-Fisher Scientific). The inactive transcription factor NFATc1 resides in the cytosol. Elevated calcium levels in the cytosol lead to the dephosphorylation of NFATc1 by calcineurin, thereby inducing its rapid translocation into the nucleus. To increase the calcium level in the cytoplasm, U2OS cells released from nocodazole washout were treated with 0.2 µM thapsigargin, an inhibitor of the sarco-endoplasmic reticulum Ca²⁺ ATPase. Nuclear import of NFATc1 was then measured in the primary nuclei and micronuclei of U2OS cells at both 6 h and 20 h after nocodazole release. Cells were subsequently fixed with 2% paraformaldehyde and import was quantified based on nuclear fluorescence. To measure nuclear import of IBB-DiHcRed⁵⁰, U2OS cells were transiently transfected with the IBB-DiHcRed-plasmid (gift from J. Ellenberg) using Lipofectamine 2000 (Invitrogen). Cells were then synchronized for 10 h with nocodazole, washed as previously described, and released into the cell cycle. Cells were fixed with 2% paraformaldehyde both 6 h and 20 h later, and import was quantified based on nuclear fluorescence.

46. Bekker-Jensen, S. *et al.* Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J. Cell Biol.* **173**, 195–206 (2006).
47. Franz, C. *et al.* MEL-28/ELYS is required for the recruitment of nucleoporins to chromatin and postmitotic nuclear pore complex assembly. *EMBO Rep.* **8**, 165–172 (2007).
48. Peterson-Roth, E., Reynolds, M., Quiervryn, G. & Zhitkovich, A. Mismatch repair proteins are activators of toxic responses to chromium-DNA damage. *Mol. Cell Biol.* **25**, 3596–3607 (2005).
49. Prasanth, S. G., Mendez, J., Prasanth, K. V. & Stillman, B. Dynamics of pre-replication complex proteins during the cell division cycle. *Phil. Trans. R. Soc. Lond. B* **359**, 7–16 (2004).
50. Dultz, E. *et al.* Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. *J. Cell Biol.* **180**, 857–865 (2008).

Creation and diagnosis of a solid-density plasma with an X-ray free-electron laser

S. M. Vinko¹, O. Ciricosta¹, B. I. Cho², K. Engelhorn², H.-K. Chung³, C. R. D. Brown⁴, T. Burian⁵, J. Chalupský⁵, R. W. Falcone^{2,6}, C. Graves⁷, V. Hájková⁵, A. Higginbotham¹, L. Juha⁵, J. Krzywinski⁷, H. J. Lee⁷, M. Messerschmidt⁷, C. D. Murphy¹, Y. Ping⁸, A. Scherz⁷, W. Schlott⁷, S. Toleikis⁹, J. J. Turner⁷, L. Vysin⁵, T. Wang⁷, B. Wu⁷, U. Zastrau¹⁰, D. Zhu⁷, R. W. Lee⁷, P. A. Heimann², B. Nagler⁷ & J. S. Wark¹

Matter with a high energy density ($>10^5$ joules per cm^3) is prevalent throughout the Universe, being present in all types of stars¹ and towards the centre of the giant planets^{2,3}; it is also relevant for inertial confinement fusion⁴. Its thermodynamic and transport properties are challenging to measure, requiring the creation of sufficiently long-lived samples at homogeneous temperatures and densities^{5,6}. With the advent of the Linac Coherent Light Source (LCLS) X-ray laser⁷, high-intensity radiation ($>10^{17}$ watts per cm^2 , previously the domain of optical lasers) can be produced at X-ray wavelengths. The interaction of single atoms with such intense X-rays has recently been investigated⁸. An understanding of the contrasting case of intense X-ray interaction with dense systems is important from a fundamental viewpoint and for applications. Here we report the experimental creation of a solid-density plasma at temperatures in excess of 10^6 kelvin on inertial-confinement time-scales using an X-ray free-electron laser. We discuss the pertinent physics of the intense X-ray-matter interactions, and illustrate the importance of electron-ion collisions. Detailed simulations of the interaction process conducted with a radiative-collisional code show good qualitative agreement with the experimental results. We obtain insights into the evolution of the charge state distribution of the system, the electron density and temperature, and the time-scales of collisional processes. Our results should inform future high-intensity X-ray experiments involving dense samples, such as X-ray diffractive imaging of biological systems, material science investigations, and the study of matter in extreme conditions.

The experiment was performed at the LCLS soft X-ray materials science instrument (SXR), where a 1.0- μm -thick aluminium (Al) foil was irradiated with 80-fs X-ray pulses at photon energies in the range 1,560–1,830 eV (at and above the Al K edge). The LCLS pulse contained $\sim 10^{12}$ photons with an energy bandwidth of $\sim 0.4\%$. The X-ray pulse was focused by means of bendable Kirkpatrick-Baez mirrors, and the focal spot was characterized *ex situ* by analysing imprints in lead tungstate (ref. 9), indicating a spot size of $9.1 \pm 0.8 \mu\text{m}^2$, that is, a peak intensity of $1.1 \times 10^{17} \text{ W cm}^{-2}$. Aluminium was studied because it is a prototypical free-electron metal, presenting all the intricacies of a high-electron-density system, yet possessing a relatively simple atomic structure. Our main target diagnostic is X-ray emission spectroscopy, sensitive to recombination from the L to the K shell, that is, to Al K α emission, in the spectral range 1,460–1,680 eV.

The absorption process in the Al foil is dominated by K-shell photoabsorption, ejecting a core electron into the continuum, because the cross-section for L-shell and valence-band photoionization is over a factor 10 smaller. After photoionization, the filling of the K-shell hole proceeds mainly by KLL Auger decay, producing an ion with two holes in the L shell. Radiative emission accounts for the remaining 3.8% of

the total recombination from the L shell¹⁰. In Fig. 1 we plot the spectrally resolved X-ray emission over a range of excitation photon energies between 1,560 and 1,830 eV. The spectra show the main K α peak around 1,487 eV, followed by a series of peaks corresponding to emission from higher charge states due to a growing number of L-shell holes.

The observation of K α emission from highly charged ions in a dense system is not of itself novel, being routinely observed when charged-particle beams (such as ions or optical-laser-generated energetic electrons) interact with solid targets^{11–17}. However, our results differ from these experiments in two important ways. First, in our experiment the K-shell holes are created exclusively by intense, quasi-monochromatic X-ray photons rather than by collisions with broadband energetic particles. The main absorption process (creating K-shell holes) will therefore only take place in a particular ion if the photon energy lies above that ion's K edge, or, given that the X-rays are so intense, if a K electron is resonantly pumped to a vacant L-shell state. It follows that the X-ray free-electron laser (FEL) acts not only as a pump—exciting electrons and heating the system—but also as a selective probe, in that

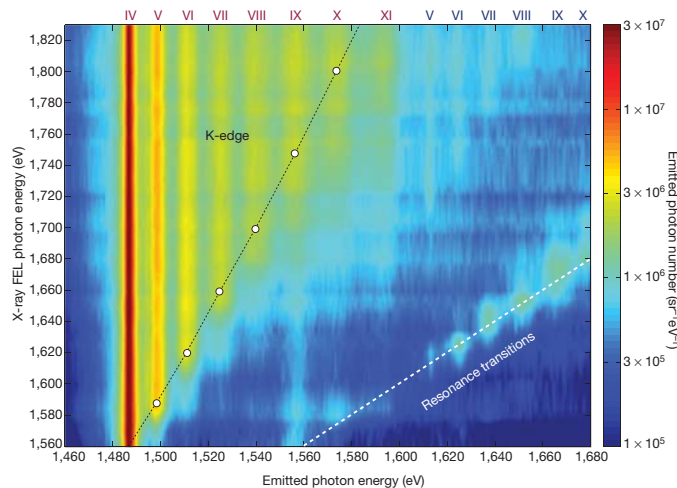


Figure 1 | Spectrally resolved K α emission as a function of the X-ray FEL excitation photon energy. The colour coding (bar on right) refers to the emission intensity on a logarithmic scale. Roman numerals (top) indicate the charge state of the emission peak: red, for states with a single K-shell hole; blue, for states with a double K-shell hole. Peaks around the resonance line (dashed white line, indicating where the FEL photon energy equals the emitted photon energy) correspond to emission from resonantly-pumped K–L transitions. Open circles, K edges for the various charge states calculated in the SCFLY code, which includes the ionization potential depression in the dense plasma according to a modified version of the Stewart-Pyatt model^{21,22}.

¹Department of Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK. ²Lawrence Berkeley National Laboratory, 1 Cyclotron Road, California 94720, USA. ³Atomic and Molecular Data Unit, Nuclear Data Section, IAEA, PO Box 100, A-1400, Vienna, Austria. ⁴Plasma Physics Department, AWE Aldermaston, Reading RG7 4PR, UK. ⁵Institute of Physics ASCR, Na Slovance 2, 18221 Prague 8, Czech Republic. ⁶Department of Physics, University of California, Berkeley, California 94720, USA. ⁷SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, California 94025, USA. ⁸Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, USA. ⁹Deutsches-Elektronensynchrotron DESY, Notkestrasse 85, 22603 Hamburg, Germany. ¹⁰IQO, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany.

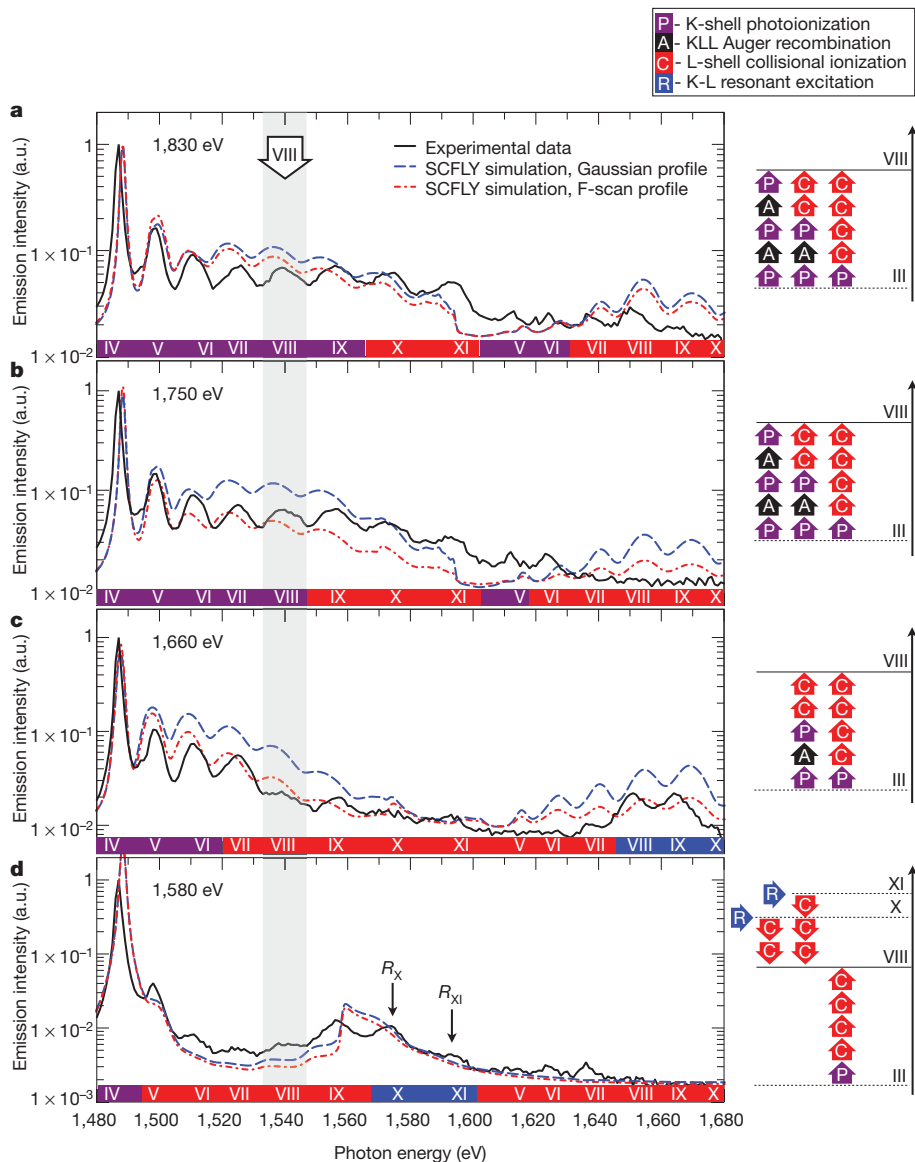


Figure 2 | Experimental and theoretical spectra, and charge-state emission pathways. Left, calculated emission spectra (dashed lines) according to the SCFLY code for four X-ray energies (**a**, 1,830 eV; **b**, 1,750 eV; **c**, 1,660 eV; **d**, 1,580 eV), compared with the experimental results presented in Fig. 1 (solid lines). The two simulation curves correspond to different X-ray intensity distributions on target: a Gaussian profile and the measured F-scan profile. Also shown (right) are the various pathways leading to emission from an ion with five L-shell holes in the final state (VIII), observed as the fifth emission peak: these pathways depend strongly on the wavelength of the X-ray excitation. R_X and R_{XI} correspond to emission from resonant transitions between K- and L-shell states in an ion containing seven or eight L-shell holes, respectively (states X and XI). Colour bars below each plot on left indicate the dominant mechanism contributing to the emission intensity for the various charge states (Roman numerals). a.u., arbitrary units.

the K-shell holes and subsequent $K\alpha$ emission only occur from a select range of ion charge states, depending on the FEL photon energy. As shown in Fig. 1, the range of ions from which we see K-shell emission is indeed strongly dependent on the energy of the incident photon: at photon energies just above the neutral K edge only a limited number of ionization stages is observed, whereas at the highest photon energy of 1,830 eV we observe all possible ionization stages corresponding to the filling of a single K-shell hole. In addition, at specific photon pump energies, we observe emission from resonantly pumped single- and double-hole K-shell states, marked as resonance transitions in Fig. 1. We label all emission peaks according to their K- and L-shell charge state (the three M-shell electrons are free-electron-like in ground-state Al): that is, the main $K\alpha$ peak corresponds to charge state IV and the higher energy peaks correspond to states in the range V–XI. The second important difference to note is that in our experiment the population of L-shell states, and the charge state distribution (CSD), is seen to be determined predominantly by electron–ion collisional processes within the heated dense system, rather than directly by the X-ray pulse. Therefore, provided the photon energy is sufficiently high to satisfy the probe condition described above for all charge states, the $K\alpha$ emission spectrum contains information on the CSD of the hot system, on the electron temperature and on electron–ion collision rates (best observed in the case of resonances where not one but several emission peaks are observed at a single pump energy).

This pivotal role of collisions in determining the observed spectra and the CSD is evinced by two main features in the data. The first is the relatively uniform intensity of the $K\alpha$ emission as a function of charge state. For isolated atoms, successive sequential photoionization of the K-shell occurs, with the dominant ionization pathway being a P–A–P–A... process (where P indicates photoionization, and A Auger decay)⁸. As each P–A sequence results in two electrons being ejected from the L shell, the populations of ions with an even number of ejected electrons dominate over those with an odd number ejected. In contrast, for the solid-density system, the dynamics are distinctly different, and we observe no odd–even asymmetry in the $K\alpha$ emission. This is due to the high electron–ion collision rates, which compete with both the P and A processes. Our simulations (below) show that the collisional and Auger rates are comparable at the peak of the pulse at the highest intensities. The second piece of evidence for the important role of collisions is the observation of emission corresponding to resonantly pumped single and double-core-hole states in ion stages for which direct photo-ionization of a K-shell electron is energetically forbidden. As the number of electrons in the L shell decreases, the reduced screening results in tighter binding of the K electrons, increasing the K-edge energy as a function of ion stage (see Fig. 1). Thus, for a particular photon energy, the ejection of a K electron is only possible up to a certain ion stage, explaining the observed fundamental trend where $K\alpha$ emission is seen from higher charge states as the photon energy is

increased. Nevertheless, this principle can be violated, as the density is so high that additional collisional ionization occurs, promoting further L electrons to the continuum (in addition to those ionized by Auger recombination). Such collisional ionization takes place for ions with two K electrons present and, importantly, is also sufficiently rapid that it can produce further L-shell vacancies in ions already containing a K-shell hole (that is, after photoabsorption, yet before Auger or radiative decay). In the former case, ions with K-edge energies higher than the incident photon energy can be created, but further photo-excitation of K-shell electrons to the continuum for these ions is no longer possible, as the K edge exceeds the photon energy. In the latter case, two interesting effects arise.

First, some $K\alpha$ emission is still seen from ion stages which have K-edge energies higher than the photon energy, via a process whereby photo-excitation of the K shell occurs at a lower ion stage, collisions ionize further L-shell electrons in the presence of the K-shell hole, then radiative decay occurs producing $K\alpha$ emission. This effect can be seen in Fig. 1 by comparing the spectrum for a given LCLS photon energy with the calculated energy of the K edge for a particular ion stage. Although accurate independent measurements of the K-edge positions in these hot-dense conditions are not available (we note that our results provide some information in this regard, albeit blurred by strong collisional effects), forcing us to rely on computational modelling, we note that in order to explain the spectra in terms of a K-edge shift alone in the absence of collisional ionization, the calculated shifts would need to be incorrect by several hundred electron volts for the highest charge states. Such a large discrepancy is implausible, but more importantly, would also lead to severely inconsistent simulation results for the emission spectrum. We further note that this observation also cannot be explained by multi-photon ionization, which scales with the square of wavelength, and is negligible even at these intensities in the X-ray regime.

Second, resonant photo-pumping of K-shell electrons to the L shell (rather than to the continuum) is energetically permitted for ions with both a single and a double K-shell hole in the final state. Radiative decay of these states produces the resonance emission highlighted in Fig. 1. Such emission appears prominently for states with two holes in the K-shell (marked as V–X for X-ray excitation in the range 1,600–1,680 eV), as well as from highly ionized states with a single K-shell hole (emission from peaks X and XI at excitation energies around 1,580 eV). In particular, we note that several peaks are observed on resonance rather than just one, indicating that the collisional rates are similar to, or larger than, Auger recombination rates.

To further interpret our results, we have performed time-dependent population kinetics calculations, using the SCFLY code¹⁸. These

simulations provide atomic level populations, CSDs, emission spectra, total absorbed energies and electron temperatures. The calculations are based on a rate-equation model that includes sequential single-photon absorption, spontaneous and stimulated emission, collisional excitation, ionization and their inverses, as well as the effects of opacity via an escape-factor formalism. We plot the calculated spectra for four X-ray excitation energies compared with our experimental measurements in Fig. 2. The interaction processes leading to the observed emission from a specific charge state (VIII) are shown alongside for the four X-ray photon energies, illustrating the different pathways resulting from a combination of direct K-shell photoionization, electron collisional ionization and resonant K–L transitions. We note that the emission intensity from the various charge states is not trivially related to the population of those charge states.

The simulated spectra successfully reproduce the experimental trends in a qualitative manner, however, quantitative differences can still be observed for several charge state emission intensities. This is in part due to uncertainties in the experimental intensity distribution on target, which is challenging to accurately account for (illustrated by the two simulation curves in Fig. 2, corresponding to different spatial profiles), in part to the simplifications made necessary for an efficient atomic-kinetics code, but mostly to uncertainties regarding fundamental physical parameters in warm and hot dense systems, such as the rates for inner-shell transitions, collisional ionization, Auger recombination in highly ionized hollow atoms, and the ionization potential depression. This last point regarding physical parameters is particularly acute in the case of emission from double-core-hole states, where the agreement between experiment and simulation is poorest. Nonetheless, given the qualitative agreement with the spectra and the straightforward energy absorption mechanism, the combined effect of these uncertainties on the basic thermodynamic quantities is not expected to be significant. The calculated electron temperature, free-electron density and CSD are shown in Fig. 3. Temperatures in excess of 100 eV are found, which do not depend strongly on the excitation wavelength. Further, higher X-ray photon energies invariably yield higher temperatures by the end of the pulse, despite having an initially lower absorption cross-section. This can be explained by thermal collisions governing the CSD: as the temperature rises, low charge states are depleted and higher-energy X-rays, which couple more efficiently to higher charge state transitions, deposit more energy in the system. To quantitatively evaluate this collisional effect, we illustrate in Fig. 4 the photoionization, Auger recombination and collisional ionization rates as a function of time for transitions populating and de-populating the single-core-hole charge state VII. We observe that by the peak of the pulse, at temperatures above 100 eV, electron–ion collisional

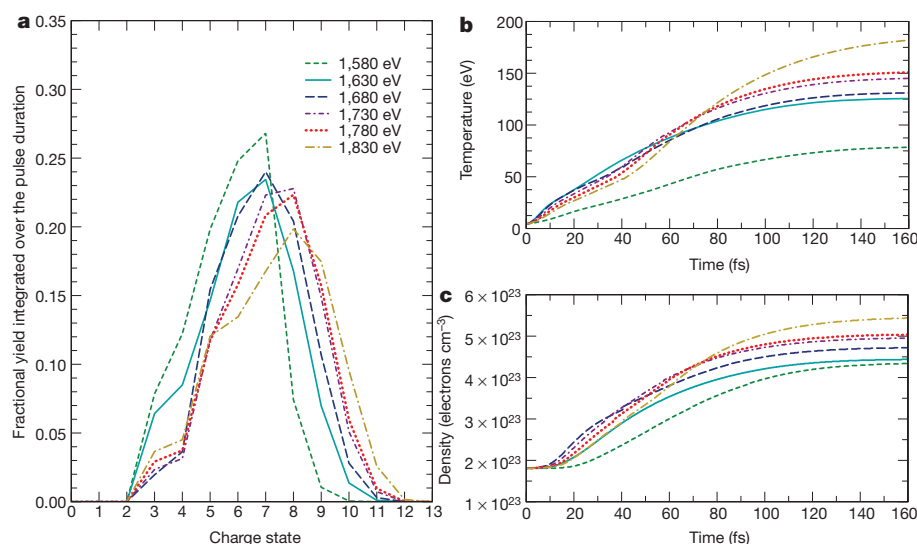


Figure 3 | Temperature, density and charge state distribution at peak intensity. **a**, Calculated charge state distribution, integrated over the duration of a temporally Gaussian pulse. The calculations correspond to the centre of the X-ray spot, at the highest spatial intensities. **b**, **c**, The temperature (**b**) and density (**c**) of the electronic system over a range of excitation energies. The peak of the X-ray pulse is centred at 80 fs, with a FWHM of 80 fs.

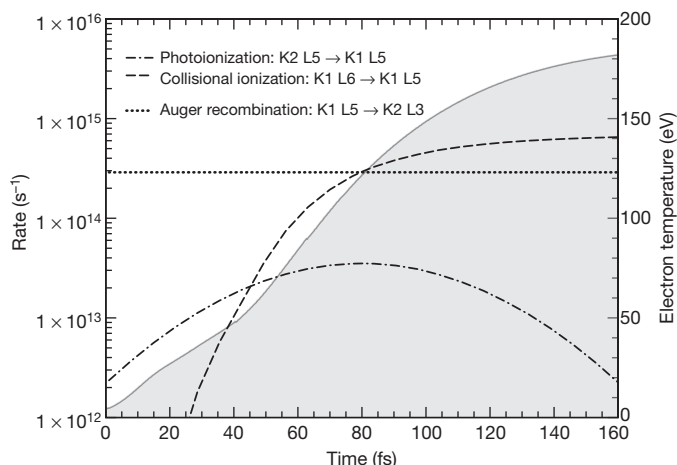


Figure 4 | Rates for atomic and collisional processes. Comparison between K-shell photoionization, collisional ionization and Auger recombination rates (broken curves, left-hand vertical axis) involved in populating and depopulating charge state VII at an X-ray photon energy of 1,830 eV. This state is one of the most populated for conditions leading to a peak temperature of ~ 100 eV. The calculation is conducted with a Gaussian beam (80 fs FWHM) centred at 80 fs. The electron temperature is given by the solid curve with grey shading under (right-hand vertical axis). By the peak of the pulse at 80 fs, the collisional ionization rate exceeds all others, dominating the population dynamics.

processes dominate the evolution of the CSD, and the dynamics of the system is determined by its temperature rather than by the intense X-ray excitation field. This result is fully consistent with our experimental results shown in Fig. 1.

Finally, in the context of creating well-defined states of matter in extreme conditions, we note that the duration of the pulse, heating of the electrons, photoionization and Auger decay are all rapid processes compared with atomic motion and hydrodynamic expansion of the target. For example, an aluminium ion at 150 eV travels only 5 nm in 100 fs, and thus a crude estimate of the hydrodynamic expansion time would be 20 ps for a 1 μm target. Given that electron-ion relaxation times are expected to be less than 10 ps (ref. 19), these results bode well for the creation of uniform-density hot and warm dense matter.

METHODS SUMMARY

A schematic illustration of the experimental set-up is given in Supplementary Fig. 1. The spectroscopy was conducted by means of a flat ammonium dihydrogen phosphate crystal (ADP) in the $\langle 101 \rangle$ orientation with a spectral resolution of 0.9–1.3 eV across the photon energy range of interest. All spectroscopic data were collected in single-shot acquisition mode, with a fresh region of the target exposed for every shot. Figure 1 was obtained by binning a total of around 550 single-shot spectra according to the measured X-ray photon energy in 5-eV bins.

Simulations of the experimental emission were conducted using an adapted version of the collisional-radiative code SCFLY¹⁸, which is based on the widely available FLYCHK simulation suite²⁰. SCFLY solves the rate equations for non-LTE (local thermal equilibrium) plasmas using a super-configuration description of the atomic structure of the system and provides a spectral module, which uses the computed level populations to calculate the X-ray emission spectra. Several important improvements have been included to make SCFLY applicable to X-ray FEL-driven systems: an increased number of super-configurations, the inclusion of hollow atoms and a self-consistent, absorption-driven temperature calculation. Each calculation provides an emission spectrum for a single X-ray intensity. To account for the experimental intensity distribution on target, 26 simulations covering the full range of probed intensities were computed for each X-ray excitation energy. The final spectra are obtained by averaging these over the three-dimensional interaction region, weighted according to the intensity distribution profile measured by the imprints technique in lead tungstate (PbWO_4). This intensity distribution, named the F-scan profile, which is obtained following the method described in ref. 9, is compared to the Gaussian intensity distribution consistent with the measured spot size (full-width at half-maximum, FWHM) in Fig. 2.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 August; accepted 29 November 2011.

Published online 25 January 2012.

1. Taylor, R. J. *The Stars: Their Structure and Evolution* 2nd edn (Cambridge Univ. Press, 1994).
2. Chabrier, G. Plasma physics and planetary astrophysics. *Plasma Phys. Contr. Fusion* **51**, 124014 (2009).
3. Helled, R., Anderson, J. D., Podolak, M. & Schubert, G. Interior models of Uranus and Neptune. *Astrophys. J.* **726**, 15 (2011).
4. Lindl, J. Development of the indirect-drive approach to inertial confinement fusion and the target physics basis for ignition and gain. *Phys. Plasmas* **2**, 3933–4024 (1995).
5. Remington, B. A., Drake, R. P. & Ryutov, D. D. Experimental astrophysics with high power lasers and Z pinches. *Rev. Mod. Phys.* **78**, 755–807 (2006).
6. Nagler, B. et al. Turning solid aluminium transparent by intense soft X-ray photoionization. *Nature Phys.* **5**, 693–696 (2009).
7. Linac Coherent Light Source. <https://lcls.slac.stanford.edu> (2011).
8. Young, L. et al. Femtosecond electronic response of atoms to ultra-intense X-rays. *Nature* **466**, 56–61 (2010).
9. Chalupský, J. et al. Spot size characterization of focused non-Gaussian X-ray laser beams. *Opt. Express* **18**, 27836–27845 (2010).
10. Bambynek, W. et al. X-Ray fluorescence yields, Auger, and Coster-Kronig transition probabilities. *Rev. Mod. Phys.* **44**, 716–813 (1972).
11. Saemann, A. et al. Isochoric heating of solid aluminum by ultrashort laser pulses focused on a tapered target. *Phys. Rev. Lett.* **82**, 4843–4846 (1999).
12. Rossall, A. K. et al. X-ray back-lighter characterization for iron opacity measurements using laser-produced aluminium K-alpha emission. *J. Phys. At. Mol. Opt. Phys.* **43**, 155403 (2010).
13. Bailey, J. et al. Observation of K α X-ray satellites from a target heated by an intense ion beam. *Laser Part. Beams* **8**, 555–562 (1990).
14. Burnett, N. H., Enright, G. D., Avery, A., Loen, A. & Kieffer, J. C. Time-resolved K α spectra in high-intensity laser-target interaction. *Phys. Rev. A* **29**, 2294–2297 (1984).
15. Back, C. A. et al. Study of K α absorption structures in a subcritical-density laser-produced plasma. *Phys. Rev. A* **46**, 3405–3412 (1992).
16. Rousse, A. et al. Efficient K α x-ray source from femtosecond laser-produced plasmas. *Phys. Rev. E* **50**, 2200–2207 (1994).
17. Gauthier, J.-C. et al. Theoretical and experimental studies of laser-produced plasmas driven by high-intensity femtosecond laser pulses. *Phys. Plasmas* **4**, 1811–1817 (1997).
18. Chung, H.-K., Chen, M. H. & Lee, R. W. Extension of atomic configuration sets of the non-LTE model in the application to the K α diagnostics of hot dense matter. *High Energy Density Phys.* **3**, 57–64 (2007).
19. Siwick, B. J., Dwyer, J. R., Jordan, R. E. & Miller, R. J. D. An atomic-level view of melting using femtosecond electron diffraction. *Science* **302**, 1382–1385 (2003).
20. Chung, H.-K., Chen, M. H., Morgan, W. L., Ralchenko, Y. & Lee, R. W. FLYCHK: Generalized population kinetics and spectral model for rapid spectroscopic analysis for all elements. *High Energy Density Phys.* **1**, 3–12 (2005).
21. Stewart, J. C. & Pyatt, K. D. Jr. Lowering of ionization potentials in plasmas. *Astrophys. J.* **144**, 1203–1211 (1966).
22. Chung, H.-K., Lee, R. W., Chen, M. H. & Ralchenko, Y. *The How To for FLYCHK* (NIST, 2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Portions of this research were carried out on the SXR instrument at the LCLS, a division of SLAC National Accelerator Laboratory and an Office of Science user facility operated by Stanford University for the US Department of Energy. The SXR instrument and the Resonant Coherent Imaging (RCI) endstation are funded by a consortium whose membership includes the LCLS, Stanford University through the Stanford Institute for Materials Energy Sciences (SIMES), Lawrence Berkeley National Laboratory (LBNL), University of Hamburg through the BMBF priority programme FSP 301, and the Center for Free Electron Laser Science (CFEL). S.M.V., O.C. and J.S.W. thank the UK EPSRC for funding (EP/F020449/1 and EP/H035877/1). B.I.C., K.E., R.W.F. and P.A.H. acknowledge US DOE Basic Energy Science contract DE-AC03-76SF00098 and SSAA programme contract DE-FG52-06NA26212. T.B., J.C., L.J. and L.V. appreciate funding by grants LC510, LC528, LA08024, ME10046, P108/11/1312, P205/11/0571, IAAX00100903 and KAN300100702. U.Z. thanks the German Ministry for Education and Research (BMBF) for funding under FSP 301. C.D.M. was supported by UK EPSRC (EP/G007187/1). We also thank G. Gregori (Oxford University) for discussions.

Author Contributions S.M.V., B.I.C., K.E., C.R.D.B., A.H., H.J.L., C.D.M., Y.P., S.T., U.Z., P.A.H., B.N. and J.S.W. performed the experiment and acquired the data. B.I.C., K.E., R.W.F. and P.A.H. analysed the data. S.M.V., O.C., H.-K.C., R.W.L. and J.S.W. performed the theoretical work. S.M.V., R.W.L. and J.S.W. wrote the paper. M.M., W.S. and J.J.T. operated the SXR beamline and the LCLS diagnostics. C.G., A.S., T.W., B.W. and D.Z. operated the RCI endstation. T.B., J.C., V.H., L.J., J.K. and L.V. performed the spot-size characterisation and analysis. All authors contributed to the work presented here and to the final paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.M.V. (sam.vinko@physics.ox.ac.uk).

METHODS

Experimental set-up. During the experiment, the total pulse energy was recorded on a shot-to-shot basis using a gas detector in the LCLS front-end enclosure. We measured an average pulse energy of 0.8–1.4 mJ, depending on the photon wavelength. Approximately 30–34% of this energy is delivered to the target owing to the transmission of the beamline optics between the gas detector and the Resonant Coherent Imaging (RCI) experimental chamber²³. During the course of the experiment, we scanned a photon energy range of 1,580–1,830 eV. The photon energy values are derived from the electron beam energies provided by the LCLS online diagnostics, the values of which were calibrated at the Al K edge both before and after the data acquisition. This method provides an accuracy of better than 5 eV in the absolute calibration, and is approximately constant over the entire range of energies investigated.

Data analysis. The noise filtering for the data analysis was conducted by analysing the number of counts per CCD (charge-coupled device) pixel: photons of 1,600 eV are seen to produce approximately 150 counts above the thermal noise level on the CCD detector, which is sufficient to isolate the signal and reduce the noise considerably. Any change in beam position on the sample will alter the crystal diffraction angle slightly, causing a shift in where wavelengths are mapped on to the detector. We have accounted for this positioning jitter effect by cross-correlating each single shot with the average spectrum and adjusting the positions of the $K\alpha$ satellite peaks accordingly to overlap. The wavelength axis of the final spectrum is then calibrated by fitting the positions of the main $K\alpha$ peak and its satellites to previously published values¹³. This eliminates the need to use geometric considerations to map the detector pixels to wavelength and is considerably more accurate. We did not measure the temporal profile of the LCLS beam or the pulse length. Instead, the LCLS electron bunch length was measured (80 fs FWHM) and assumed to be a good indication of the X-ray pulse duration. However, this value should only be considered as an upper bound on the pulse duration, with recent experiments reporting results consistent with 2–3 times shorter pulses^{8,24}.

Modelling. Simulations of the experimental emission were conducted using the SCFLY code^{18,20} adapted to simulate FEL-driven systems. It has recently been reported that these modifications enable SCFLY to successfully simulate X-ray–matter interactions in the non-collisional regime limit²⁵. For the work presented here, which extends this study to the highly collisional regime, a complete set of super-configurations for all ionic species of aluminium up to the $n = 3$ shell are considered in the code. We have further verified that the spectrum exhibits a negligible dependence on the temporal structure of the X-ray pulse, provided the total fluence is constant, by comparing calculated spectra obtained by using a Gaussian temporal profile with a simulated self-amplified spontaneous emission FEL pulse of equal duration and fluence, following ref. 26.

The energies used in the simulations are calculated according to the relativistic configuration average model following a Dirac–Hartree–Slater method²⁷. The agreement in peak position that can be observed between simulation and experiment, within ~ 3 eV (0.2%) for the highest charge states, was deemed sufficient within the context of our investigation. It can however be improved on by using a fine structure model for the energy levels, albeit at a significant additional computational cost.

23. Sorokin, A. A., Jastrow, U., Juranić, P., Kapitzki, S. & Tiedtke, K. *Report on Pulse Energy Monitoring at the SXR Beamline Using Gas-Monitor Detectors* (Technical report, SLAC, 2010).
24. Düsterer, S. *et al.* Femtosecond x-ray pulse length characterization at the Linac Coherent Light Source free-electron laser. *New J. Phys.* **13**, 093024 (2011).
25. Cericosta, O., Chung, H.-K., Lee, R. W. & Wark, J. S. Simulations of neon irradiated by intense X-ray laser radiation. *High Energy Density Phys.* **7**, 111–116 (2011).
26. Bonifacio, R., De Salvo, L., Pierini, P., Piovella, N. & Pellegrini, C. Spectrum, temporal structure, and fluctuations in a high-gain free-electron laser starting from noise. *Phys. Rev. Lett.* **73**, 70–73 (1994).
27. Chen, M. H., Crasemann, B. & Mark, H. Relativistic K-shell Auger rates, level widths, and fluorescence yields. *Phys. Rev. A* **21**, 436–441 (1980).

Quantum-coherent coupling of a mechanical oscillator to an optical cavity mode

E. Verhagen^{1*}, S. Deléglise^{1*}, S. Weis^{1,2*}, A. Schliesser^{1,2*} & T. J. Kippenberg^{1,2}

Optical laser fields have been widely used to achieve quantum control over the motional and internal degrees of freedom of atoms and ions^{1,2}, molecules and atomic gases. A route to controlling the quantum states of macroscopic mechanical oscillators in a similar fashion is to exploit the parametric coupling between optical and mechanical degrees of freedom through radiation pressure in suitably engineered optical cavities^{3–6}. If the optomechanical coupling is ‘quantum coherent’—that is, if the coherent coupling rate exceeds both the optical and the mechanical decoherence rate—quantum states are transferred from the optical field to the mechanical oscillator and vice versa. This transfer allows control of the mechanical oscillator state using the wide range of available quantum optical techniques. So far, however, quantum-coherent coupling of micromechanical oscillators has only been achieved using microwave fields at millikelvin temperatures^{7,8}. Optical experiments have not attained this regime owing to the large mechanical decoherence rates⁹ and the difficulty of overcoming optical dissipation¹⁰. Here we achieve quantum-coherent coupling between optical photons and a micromechanical oscillator. Simultaneously, coupling to the cold photon bath cools the mechanical oscillator to an average occupancy of 1.7 ± 0.1 motional quanta. Excitation with weak classical light pulses reveals the exchange of energy between the optical light field and the micromechanical oscillator in the time domain at the level of less than one quantum on average. This optomechanical system establishes an efficient quantum interface between mechanical oscillators and optical photons, which can provide decoherence-free transport of quantum states through optical fibres. Our results offer a route towards the use of mechanical oscillators as quantum transducers or in microwave-to-optical quantum links^{11–15}.

Mechanical oscillators are at the heart of many precision experiments and can exhibit exceptionally low dissipation. The possibility of controlling the quantum states of engineered micro- or nano-mechanical oscillators has been a subject of long-standing interest^{3,16}. Recent experiments have successfully prepared such devices in their quantum ground state, either using standard cryogenic techniques⁷ or via dynamical back-action cooling^{8,10}. Full control over the quantum state of a mechanical oscillator can be achieved by coupling it to an auxiliary system—whose quantum state can be controlled and measured—under the condition that the coherent coupling rate exceeds the decoherence rate of each of the subsystems. Such quantum-coherent coupling and control at the single-phonon level has been recently demonstrated in the microwave domain by achieving strong coupling of a gigahertz piezoelectrical dilatation oscillator to a (frequency-degenerate) superconducting qubit⁷. Moreover, an electromechanical system has reached the quantum-coherent regime using parametric coupling of a megahertz micromechanical oscillator to a gigahertz microwave resonator^{8,17}. Achieving quantum-coherent coupling in the optical domain would provide an interface between mechanical quantum states and optical fields, whose low thermal occupancy at room temperature allows decoherence-free propagation of quantum states out

of a cryogenic environment and which can be coupled to a variety of quantum systems. Additionally, optical techniques benefit from the wide availability of quantum-limited detection techniques. In fact, simultaneous achievement of quantum-coherent coupling in the optical and microwave domain provides the possibility of realizing optical-to-microwave conversion of quantum states¹⁵.

Parametric optomechanical coupling occurs in an optical microcavity that simultaneously exhibits a mechanical resonator mode, whose displacement alters the optical resonance frequency. The coupling can be described by the interaction Hamiltonian $H = \hbar g_0 \hat{a}^\dagger \hat{a} (\hat{b}^\dagger + \hat{b})$, where \hat{a} (\hat{a}^\dagger) and \hat{b} (\hat{b}^\dagger) are the photon and phonon annihilation (creation) operators, respectively, \hbar is the reduced Planck constant and g_0 is the vacuum optomechanical coupling rate. In the resolved sideband regime (where the mechanical resonance frequency Ω_m exceeds the cavity energy decay rate κ), with an intense laser tuned close to the lower optomechanical sideband, one obtains in the rotating wave approximation the effective Hamiltonian

$$H = \hbar g (\hat{a} \hat{b}^\dagger + \hat{a}^\dagger \hat{b}) \quad (1)$$

for the operators \hat{a} and \hat{b} now displaced by their steady state values. We have introduced here the field-enhanced coupling rate^{18–20} $g = \sqrt{\bar{n}_c} g_0$, where \bar{n}_c denotes the average number of photons in the cavity. In the absence of decoherence, the unitary evolution given by equation (1) corresponds to swapping of the (displaced) optical and mechanical quantum states with a period of $2\pi/\Omega_c$, where $\Omega_c = 2g$ is the coherent coupling rate, that is, the rate at which the two systems exchange energy. This state swapping is at the heart of most quantum control protocols^{11–14,21,22}. In practice, however, this unitary evolution is compromised by the coupling of both degrees of freedom to their respective environments. Hence, it is important for Ω_c to exceed both the optical decoherence rate κ and the relevant mechanical decoherence rate γ . If the mechanical linewidth is limited by dissipation (as is the case in our system²³), the mechanical decoherence is caused by energy relaxation to the hot environment. Thus, both decoherence rates are defined here as the inverse of the time needed for a single excitation to be lost into the environment. Importantly, the mechanical mode is coupled to an environment that generally has a large thermal occupancy $\bar{n}_m \approx k_B T / \hbar \Omega_m$ (where k_B is Boltzmann’s constant and T is temperature) owing to the relatively small mechanical frequency. Therefore, the mechanical decoherence rate $\gamma = \Gamma_m (\bar{n}_m + 1) \approx k_B T / \hbar Q_m$ (where $Q_m = \Omega_m / \Gamma_m$; refs 1, 14) is much larger than the dissipation rate Γ_m . This means that the strong coupling regime $\Omega_c > (\kappa, \Gamma_m)$ (ref. 9) does not suffice to enable quantum control, in contrast to the situation in atomic cavity quantum electrodynamics² or in resonant coupling to superconducting qubits of sufficiently high frequency⁷. Instead, for parametrically coupled systems the more stringent condition $\Omega_c > (\kappa, \gamma)$ signals the relevant regime, which we call quantum-coherent coupling.

Here we realize quantum-coherent coupling of an optical cavity field to a micromechanical oscillator. In this regime, the system is

¹Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. ²Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany.

*These authors contributed equally to this work.

appropriately described as an optomechanical polariton whose decoherence time exceeds the period of the Rabi oscillations between light and mechanics. Although the ground state has recently been reported with optical cooling¹⁰, in that work Ω_c was much smaller than κ . In contrast, the present system achieves a coupling rate exceeding both the optical and mechanical decoherence rates, thereby satisfying the necessary conditions for full control of the quantum state of a mechanical oscillator with optical fields^{9,12–14,21,22}. The experimental setting is a micro-optomechanical system in the form of a spoke-anchored toroidal optical microcavity²⁴. Such devices exhibit whispering gallery mode resonances of high quality factor (with a typical cavity decay rate $\kappa/2\pi < 10$ MHz) coupled to mechanical radial breathing modes via radiation pressure²⁵. The vacuum optomechanical coupling rate $g_0 = (\omega_c/R)x_{\text{ZPM}}$ can be increased by reducing the radius R of the cavity (here ω_c is the optical cavity resonance frequency and x_{ZPM} is the zero point motion). However, the larger per photon force $\hbar\omega_c/R$ is then usually partially compensated by the increase in the mechanical resonance frequency Ω_m —and correspondingly smaller zero point motion, given by $x_{\text{ZPM}} = \sqrt{\hbar/(2m_{\text{eff}}\Omega_m)}$ (where m_{eff} is the effective oscillator mass). Moreover, small structures also generally feature larger dissipation through clamping losses. To compensate these opposing effects, we use an optimized spoke-anchor design (see Fig. 1 and Supplementary Information) that maintains low clamping losses and a moderate mechanical resonance frequency while reducing the dimensions of the structure. Devices fabricated in this manner (with $R = 15\ \mu\text{m}$) exhibited coupling rates as high as $g_0 = 2\pi \times 3.4$ kHz for a resonance frequency of 78 MHz and a critically coupled sideband factor $\Omega_m/\kappa = 11$.

To reduce the mechanical decoherence rate $\gamma \approx \Gamma_m \bar{n}_m$, the microcavity is embedded in a ³He cryostat (minimum temperature $T_{\text{min}} = 650$ mK)²⁶. A continuous-wave Ti:sapphire laser beam, whose phase and amplitude quadrature noises are quantum-limited for the Fourier frequencies of interest, is coupled to the microcavity using a tapered optical fibre. The weak mechanical displacement fluctuations are recorded by measuring the phase fluctuations imprinted on the field emerging from the cryostat using balanced homodyne detection. Whereas the coherent coupling rate Ω_c can be determined unambiguously by probing the coherent response of the system²⁷, the mechanical decoherence rate is affected in a non-trivial way by the light-absorption-dependent sample temperature and the mechanical mode's coupling to its environment, which is dominated by two-level fluctuators at cryogenic temperatures^{23,26}. In order to systematically assess the aforementioned effects on the decoherence rate, the coupling laser's frequency $\omega_l = \omega_c + \Delta$ (where Δ denotes the laser detuning) is varied in the vicinity of the lower mechanical sideband, while keeping

the launched power constant. This allows the displaced cavity mode \hat{a} (of frequency $|\Delta|$) to be brought in and out of resonance with the mechanical mode \hat{b} (of frequency Ω_m). For each detuning point, we acquire the coherent response of the system to an optical excitation of swept frequency $\omega_l + \Omega_{\text{mod}}$ in a first step (Ω_{mod} is the frequency difference from the coupling laser). These spectra (Fig. 2a) allow us to determine all parameters of the model characterizing the optomechanical interaction (Supplementary Information). For large detunings $|\Delta| > \Omega_m$, they essentially feature a Lorentzian response of width κ and centre frequency $|\Delta|$. The sharp dip at $\Omega_{\text{mod}} \approx \Omega_m$ originates from optomechanically induced transparency²⁷ (OMIT), and for $\Omega_m = -\Delta$, its width is approximately Ω_c^2/κ . The coupling rate, as derived from a fit of the coherent response for a laser power of 0.56 mW, is $\Omega_c = 2\pi \times (3.7 \pm 0.05)$ MHz (corresponding to an intracavity photon number of $\bar{n}_c = 3 \times 10^5$).

Additionally, for each value of the detuning, the noise spectrum of the homodyne signal is recorded in the absence of any external excitation (Fig. 2b). The observed peak represents the phase fluctuations imprinted on the transmitted light by the mechanical mode's thermal motion. The constant noise background on these spectra is the shot-noise level for the (constant) laser power used throughout the laser sweep (see Supplementary Information for details). Importantly, the amplitude of the peak is determined by the coupling to and the temperature of the environment, and therefore allows us to extract the mechanical decoherence rate. All parameters now having been measured, it is moreover possible to retrieve the mechanical displacement spectrum (Fig. 2c inset). As can be seen, for detunings close to the sideband, when the (displaced) optical and mechanical modes are degenerate, the fluctuations are strongly reduced. This effect of optomechanical resolved sideband cooling²⁸ can be understood in a simple picture: in the regime $\Omega_c \ll \kappa$, the optical decay is faster than the swapping between the vacuum in the displaced optical field and the thermal state in the mechanical oscillator. In this case, the mechanical oscillator is coupled to an effective optical bath at near-zero thermal occupancy \bar{n}_{min} with the rate $\Gamma_{\text{cool}} = \Omega_c^2/\kappa$. Ideally, $\bar{n}_{\text{min}} = \kappa^2/16\Omega_m^2 \ll 1$ is governed by non-resonant Stokes terms $\hat{a}^\dagger \hat{b}^\dagger + \hat{a} \hat{b}$ neglected in the Hamiltonian (equation (1))^{18,19}.

Evaluating the mechanical decoherence rate for $\Delta = -\Omega_m$ at a cryostat set point of 0.65 K, we find $\gamma = 2\pi \times (2.2 \pm 0.2)$ MHz—significantly smaller than Ω_c . Simultaneously, the average occupancy of the mechanical mode is reduced to $\bar{n} = 1.7 \pm 0.1$ (corresponding to $37 \pm 4\%$ ground state occupation), which is limited by the onset of normal mode splitting. Indeed, as Ω_c approaches κ , the thermal fluctuations are only partially dissipated into the optical bath, and are partially written back onto the mechanics after one Rabi cycle.

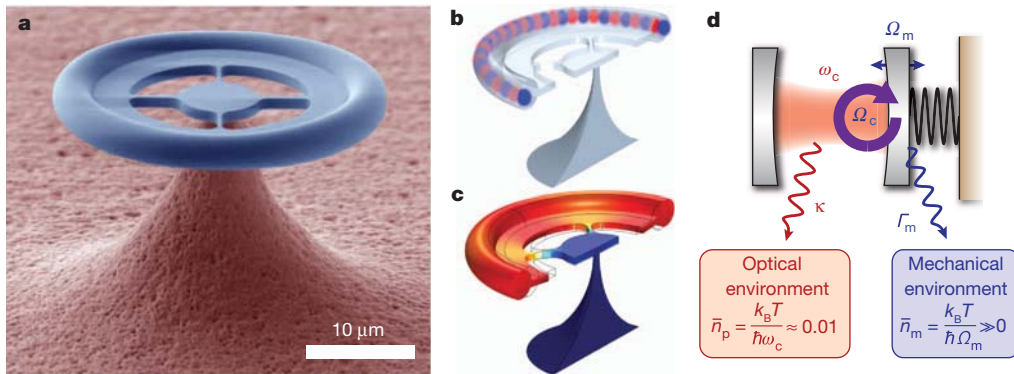


Figure 1 | Optomechanical microresonators. **a**, False-colour scanning electron micrograph of a spoke-anchored toroidal resonator 31 μm in diameter used for the optomechanical experiments reported in this work. **b**, Sketch of an optical whispering gallery mode in the microresonator (colours indicate optical phase). **c**, Simulated displacement (exaggerated for clarity) of the fundamental radial breathing mode of the structure. **d**, Equivalent optomechanical Fabry–

Pérot cavity: quantum-coherent coupling is achieved when the enhanced coupling rate Ω_c is comparable to or exceeds the optical and mechanical decoherence rates ($\kappa, \Gamma_m \bar{n}_m$). Owing to the large asymmetry between mechanical and optical frequencies, the occupancies of the two environments are widely different.

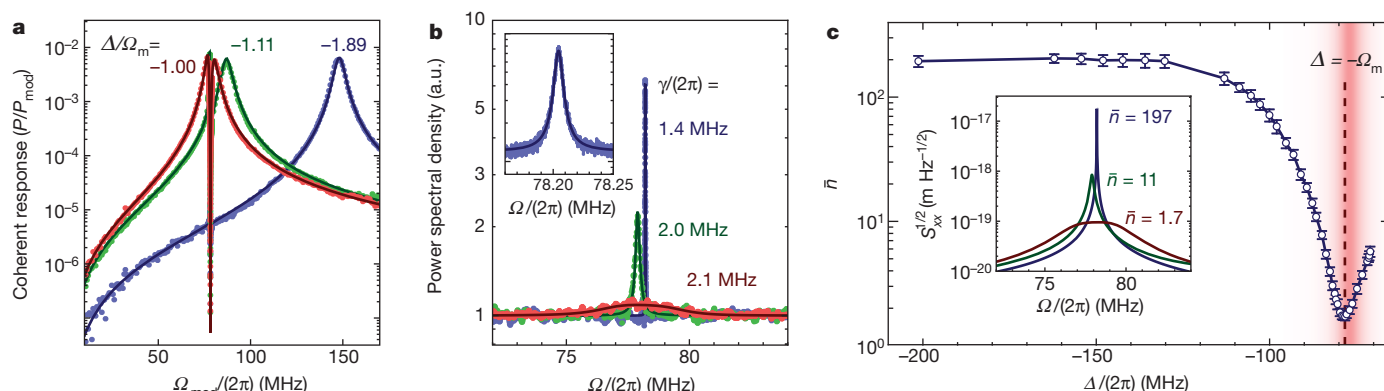


Figure 2 | Optomechanical interaction in the weak coupling regime ($\Omega_c \lesssim \kappa/2$). **a**, The coherent response of the system is obtained by sweeping a weak probe beam (at frequency $\omega_l + \Omega_{\text{mod}}$) over the cavity resonance and recording the homodyne signal P (normalized to the power P_{mod} used to create the probe beam). The fitted detuning (Δ/Ω_m) is indicated for each of the traces. The fit provides accurate estimation of the coupling rate Ω_c via the OMIT window (Supplementary Information). **b**, The measured Brownian noise spectrum in the absence of a coherent probe for each of the detunings in **a**. The spectra are corrected for the detector response and a small contribution of Guided Acoustic Wave Brillouin scattering (GAWBS) in the fibres (see

Supplementary Information). The inset shows a close-up of the spectrum obtained for $\Delta/\Omega_m = -1.89$. The indicated decoherence rates γ are deduced from the amplitude of each noise spectrum (curves show the result of the model with fitted γ). **c**, Retrieved occupancy as a function of detuning. The minimum occupancy is 1.7 ± 0.1 , meaning that the oscillator has a $37 \pm 4\%$ probability of occupying its ground state. The inset shows the inferred mechanical displacement spectral density S_{xx} for detunings used in **a** and **b** calculated using the extracted parameters. Error bars, estimated s.d. (see Supplementary Information).

We subsequently increase the strength of the coupling field to reach $\Omega_c \approx \kappa$. The signature of normal mode splitting^{9,17,19,20} can be seen from both the coherent response and the fluctuation spectra (Fig. 3). Both detuning series exhibit a clear anti-crossing, the splitting frequency Ω_c being 5.7 MHz. Note that, in contrast to microwave experiments⁸, no squashing is observed in the normal mode splitting data. The decoherence rate (Fig. 3d) is slightly raised compared to Fig. 2 owing to laser heating and a higher buffer gas temperature of 0.8 K, amounting to $\gamma = 2\pi \times (5.6 \pm 0.9)$ MHz at the lower mechanical sideband. We hence demonstrate $\Omega_c/\gamma = 1.0$, which constitutes an improvement of four orders of magnitude over the only previous work demonstrating strong coupling in the optical domain⁹, and brings the

system into the regime of quantum-coherent coupling. Even though all observed signatures are consistent with classical behaviour, entering this regime opens the door to preparation and control of non-classical states of mechanical motion with light.

As a proof of principle and as a first classical illustration of the potential of time-domain experiments, we demonstrate the dynamical exchange of a pulsed coherent excitation between the optical and mechanical degrees of freedom. By measuring the homodyne signal upon launching a weak pulse resonant with the cavity in the presence of the detuned coupling field (Fig. 4a), it is possible to directly observe the coherent exchange of energy. In the regime of weak coupling, the optical pulse excites the mechanical mode to a finite oscillation amplitude,

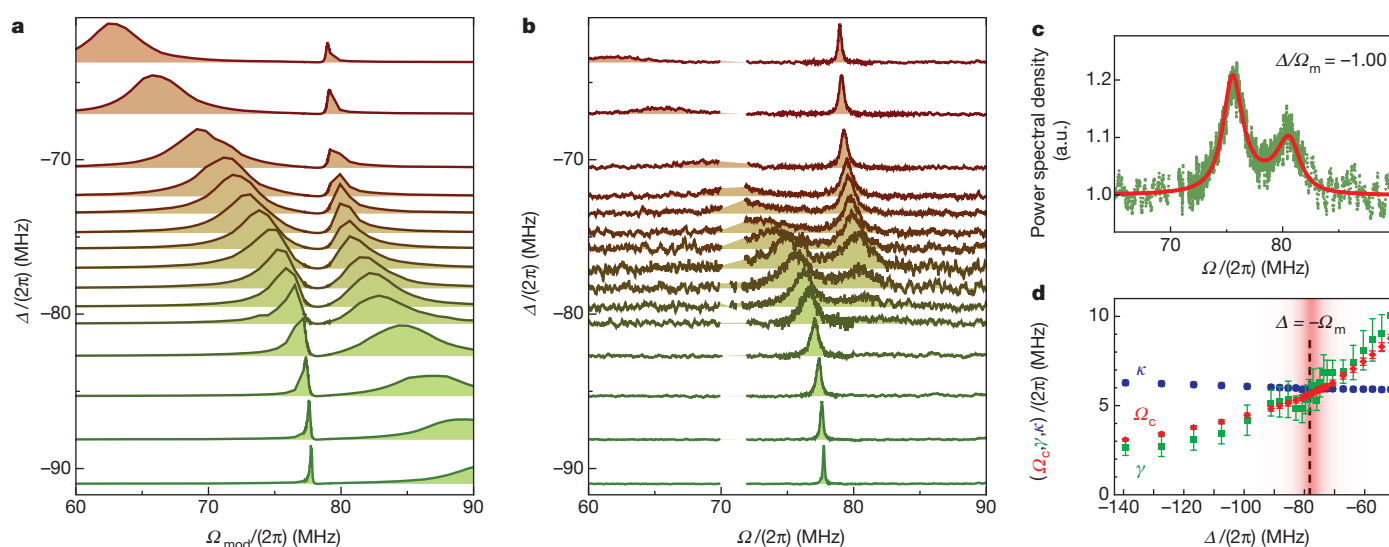


Figure 3 | Quantum-coherent coupling. **a**, **b**, The coherent optical response (**a**) and the incoherent mechanical noise spectrum (**b**) for various detunings of the coupling laser (with constant power $P = 1.4$ mW). All curves are normalized, and vertically displaced by the detuning. Evident in both panels is the avoided crossing which originates from optomechanical normal mode splitting. A second much more weakly coupled mechanical mode at 71 MHz is omitted and the curves are shaded for clarity. **c**, Homodyne noise spectrum obtained for $\Delta = -\Omega_m$. The red line is a fit of the model. Only the decoherence

rate (determining the amplitude) is fitted, while the shape is fixed through the parameters determined from the coherent response. **d**, Comparison of the mechanical (green) and optical (blue) decoherence rates with the coherent coupling rate (red) as a function of detuning. The increase of γ close to resonance reflects heating of the cavity owing to the larger amount of absorbed light. On the lower mechanical sideband, where the interaction is resonant, the decoherence rates are comparable to the coupling rate, achieving quantum-coherent coupling. Error bars, estimated s.d. (see Supplementary Information).

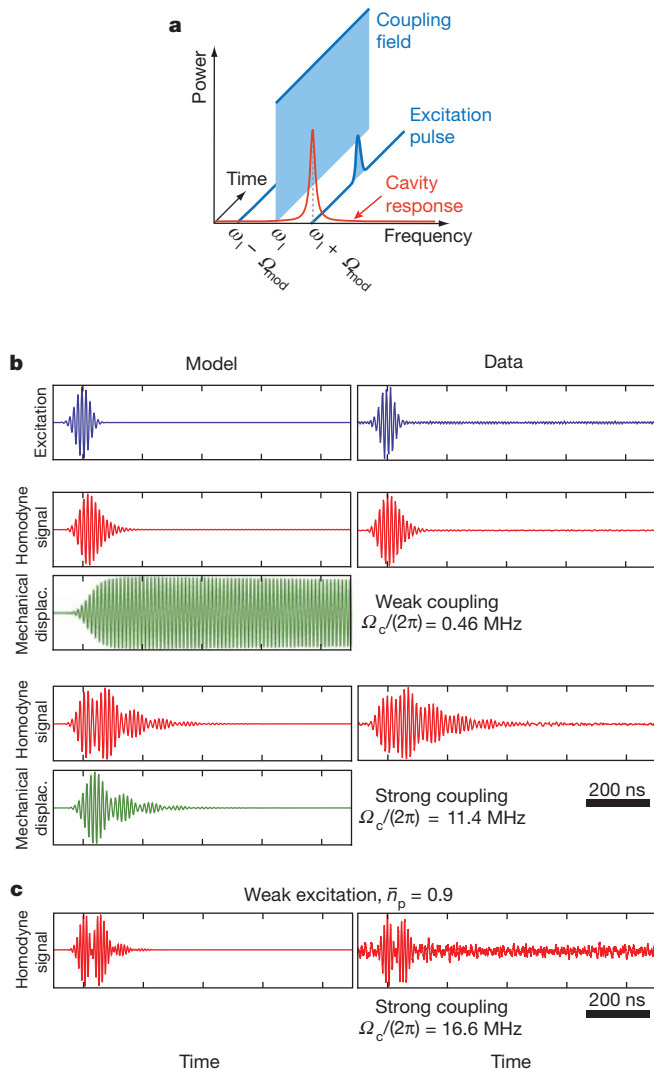


Figure 4 | Coherent exchange between the optical field and the micromechanical oscillator. This exchange is probed in the time domain as measured ('Data') and as calculated numerically ('Model'). A modulation pulse (blue traces) applied to the phase modulator creates an excitation pulse probing the dynamics of the optomechanical system in the presence of the coupling field. The response of the system is encoded into the homodyne output signal (red traces). Using the full model of the system, the mechanical displacement can be simulated in addition (green traces). **a**, Schematic of the excitation scheme, with the cavity frequency response sketched as the red curve. **b**, In the regime of weak coupling (upper panels), the optical output pulse exhibits only a weak signature of the mechanical ringdown excited by the short burst of radiation pressure. In the case of strong coupling (lower panels) the modulated envelopes of the time-domain response indicate several cycles of oscillation between coherent optical and mechanical excitations of the system. **c**, Results for weak excitation (average photon number $\bar{n}_p = 0.9$) and strong coupling.

which decays slowly at the optically damped mechanical dissipation rate. This in turn only imprints a weak signature on the homodyne signal (Fig. 4b). Increasing the laser power, we reach $\Omega_c/2\pi = 11.4$ MHz $>$ $\kappa/2\pi = 7.1$ MHz, and the envelope of the homodyne signal—corresponding to the amplitude of the measured sideband field—undergoes several cycles of energy exchange with the mechanical oscillator, before it decays with a modified rate of $(\kappa + \Gamma_m)/2 \approx \kappa/2$, corresponding to the decay rate of the excited optomechanical polariton. Using our model (which has been matched to separately taken coherent response measurements), we can derive not only the homodyne signature—which reproduces our data very well—but also the expected mechanical oscillations resulting from this pulsed excitation. These reveal how the excitation cycles continuously between the

optical and mechanical modes. Reducing the excitation further, so that one pulse contains less than one photon on average, a clear signature of several swapping cycles can still be observed (Fig. 4c). Although the average evolution observed here follows classical dynamics and is not affected by decoherence, replacing the weak coherent state in this experiment by a single photon^{13,22} is expected to result in Rabi-like oscillation of the Fock state from optics to mechanics as the system resides in the quantum-coherent regime (which we have proved above on the basis of measurements of classical signatures). A repeated quadrature measurement yielding a bimodal distribution would then provide an unambiguous signature of the quantum nature of the state after the full swap²⁹.

Obtaining quantum-coherent coupling $\Omega_c \gtrsim (\gamma, \kappa)$ has several interesting consequences. It allows the mapping of elemental quantum states of the optical field onto the mechanical mode via the use of a time-dependent coupling field. For example, preparation of the mode in the ground state or the Fock state $|1\rangle$ can be efficiently achieved in this regime using a π -pulse which swaps the thermal state of the oscillator and the quantum state of the displaced optical field. Note that the manipulation of large quantum states becomes increasingly challenging because the lifetime of the number state $|n\rangle$ scales with $1/n$. In this context, it will be beneficial to reduce spurious laser heating and employ materials with low intrinsic loss, as well as to increase the optomechanical coupling rate by further miniaturization. The regime of quantum-coherent coupling demonstrated here has been proposed as a general quantum link between electromagnetic fields of vastly different frequencies—for example, different wavelengths in the optical spectrum or microwave and optical photons^{12,15}. Electrical actuation of whispering gallery mode resonators has recently been demonstrated at an elementary level³⁰. In this context, the efficient coupling of the demonstrated system to a low-loss single mode optical fibre is beneficial. Moreover, quantum-coherent coupling enables the use of the mechanical oscillator as a transducer to link otherwise incompatible elements in hybrid quantum systems, such as solid-state spin, charge, or superconducting qubits and propagating optical fields¹⁴.

The reported experiments—which achieve quantum-coherent coupling between a micromechanical oscillator and an optical mode—represent an important step into the experimental investigation and optical quantum control of the most tangible harmonic oscillator: a mechanical vibration.

METHODS SUMMARY

The quoted value of g_0 is determined independently at room temperature via a calibrated phase modulation technique. The microcavity is evanescently coupled to a tapered optical fibre in a ^3He cryostat using piezoelectric positioning stages. For low optical power, the ^3He buffer gas enables thermalization of the resonator over the entire cryogenic temperature range in spite of its weak thermal anchoring to the substrate. A balanced homodyne detection scheme is employed, and a continuous wave Ti:sapphire laser is used to derive both the coupling field and the homodyne local oscillator. Care is taken to avoid excess noise in the coupling beam, as this would cause an effective increase of \bar{n}_{min} and preclude any quantum state manipulation. In practice it has proven crucial to eliminate phase noise originating from guided acoustic wave Brillouin scattering in the optical fibres by engineering their acoustic modes using HF etching (Supplementary Information). For the time-domain experiments, the coupling laser's phase is modulated at the mechanical resonance frequency, with a Gaussian envelope of 54 ns duration, creating a pair of sidebands—one of which is resonant with the optical cavity—that contain on average ten quanta per pulse (Supplementary Information). Although the detection used is far from optimized for time-domain experiments, a signal-to-noise ratio of 40 is achieved in Fig. 4b by averaging 250,000 traces within two minutes.

Received 5 September; accepted 9 December 2011.

1. Wineland, D. J. *et al.* Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl. Inst. Stand. Technol.* **103**, 259–328 (1998).
2. Kimble, H. J. Strong interactions of single atoms and photons in cavity QED. *Phys. Scripta* **T76**, 127–137 (1998).

3. Braginsky, V. B. *Measurement of Weak Forces in Physics Experiments* (Univ. Chicago Press, 1977).
4. Kippenberg, T. J. & Vahala, K. J. Cavity optomechanics: back-action at the mesoscale. *Science* **321**, 1172–1176 (2008).
5. Favero, I. & Karrai, K. Optomechanics of deformable optical cavities. *Nature Photon.* **3**, 201–205 (2009).
6. Clerk, A. A., Devoret, M. H., Girvin, S. M., Marquardt, F. & Schoelkopf, R. J. Introduction to quantum noise, measurement, and amplification. *Rev. Mod. Phys.* **82**, 1155–1208 (2010).
7. O'Connell, A. D. *et al.* Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).
8. Teufel, J. D. *et al.* Sideband cooling of micromechanical motion to the quantum ground state. *Nature* **475**, 359–363 (2011).
9. Gröblacher, S., Hammerer, K., Vanner, M. R. & Aspelmeyer, M. Observation of strong coupling between a micromechanical resonator and an optical cavity field. *Nature* **460**, 724–727 (2009).
10. Chan, J. *et al.* Laser cooling of a nanomechanical oscillator into its quantum ground state. *Nature* **478**, 89–92 (2011).
11. Zhang, J., Peng, K. & Braunstein, S. L. Quantum-state transfer from light to macroscopic oscillators. *Phys. Rev. A* **68**, 013808 (2003).
12. Tian, L. & Wang, H. Optical wavelength conversion of quantum states with optomechanics. *Phys. Rev. A* **82**, 053806 (2010).
13. Akram, U., Kiesel, N., Aspelmeyer, M. & Milburn, G. J. Single-photon optomechanics in the strong coupling regime. *N. J. Phys.* **12**, 083030 (2010).
14. Stannigel, K., Rabl, P., Sørensen, A. S., Zoller, P. & Lukin, M. D. Optomechanical transducers for long-distance quantum communication. *Phys. Rev. Lett.* **105**, 220501 (2010).
15. Regal, C. A. & Lehnert, K. W. From cavity electromechanics to cavity optomechanics. *J. Phys. Conf. Ser.* **264**, 012025 (2011).
16. Schwab, K. C. & Roukes, M. L. Putting mechanics into quantum mechanics. *Phys. Today* **58**, 36–42 (2005).
17. Teufel, J. D. *et al.* Circuit cavity electromechanics in the strong coupling regime. *Nature* **471**, 204–208 (2011).
18. Wilson-Rae, I., Nooshi, N., Zwerger, W. & Kippenberg, T. J. Theory of ground state cooling of a mechanical oscillator using dynamical backaction. *Phys. Rev. Lett.* **99**, 093901 (2007).
19. Marquardt, F., Chen, J. P., Clerk, A. A. & Girvin, S. M. Quantum theory of cavity-assisted sideband cooling of mechanical motion. *Phys. Rev. Lett.* **99**, 093902 (2007).
20. Dobrindt, J. M., Wilson-Rae, I. & Kippenberg, T. J. Parametric normal-mode splitting in cavity optomechanics. *Phys. Rev. Lett.* **101**, 263602 (2008).
21. Khalili, F. *et al.* Preparing a mechanical oscillator in non-Gaussian quantum states. *Phys. Rev. Lett.* **105**, 070403 (2010).
22. Romero-Isart, O. *et al.* Optically levitating dielectrics in the quantum regime: theory and protocols. *Phys. Rev. A* **83**, 013803 (2011).
23. Arcizet, O., Rivière, R., Schliesser, A., Anetsberger, G. & Kippenberg, T. J. Cryogenic properties of optomechanical silica microcavities. *Phys. Rev. A* **80**, 021803(R) (2009).
24. Anetsberger, G., Rivière, R., Schliesser, A., Arcizet, O. & Kippenberg, T. J. Ultralow-dissipation optomechanical resonators on a chip. *Nature Photon.* **2**, 627–633 (2008).
25. Kippenberg, T. J., Rokhsari, H., Carmon, T., Scherer, A. & Vahala, K. J. Analysis of radiation-pressure induced mechanical oscillation of an optical microcavity. *Phys. Rev. Lett.* **95**, 033901 (2005).
26. Rivière, R. *et al.* Optomechanical sideband cooling of a micromechanical oscillator close to the quantum ground state. *Phys. Rev. A* **83**, 063835 (2011).
27. Weis, S. *et al.* Optomechanically induced transparency. *Science* **330**, 1520–1523 (2008).
28. Schliesser, A., Rivière, R., Anetsberger, G., Arcizet, O. & Kippenberg, T. Resolved-sideband cooling of a micromechanical oscillator. *Nature Phys.* **4**, 415–419 (2008).
29. Lvovsky, A. I. *et al.* Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87**, 050402 (2001).
30. Lee, K. H., McRae, T. G., Harris, G. I., Knittel, J. & Bowen, W. P. Cooling and control of a cavity optoelectromechanical system. *Phys. Rev. Lett.* **104**, 123604 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge R. Rivière for early contributions to this project and E. Gavartin and I. Vázquez García for assistance in the early microfabrication phase. This work was supported by an ERC Starting Grant (SiMP), the DARPA/MTO ORCHID program through a grant from the AFOSR, the NCCR Quantum Science and Technology and the Swiss National Science Foundation. E.V. acknowledges a Rubicon Grant from the Netherlands Organization for Scientific Research (NWO), co-financed by a Marie Curie Cofund Action. S.D. is supported by a Marie Curie Individual Fellowship.

Author Contributions The cryogenic cooling, measurement, and sample design and fabrication was carried out jointly by S.D., E.V. and S.W. The theoretical model for data analysis was developed by A.S. All authors discussed the experimental data and jointly wrote the manuscript. T.J.K. supervised the work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.J.K. (tobias.kippenberg@epfl.ch).

Direct frequency comb spectroscopy in the extreme ultraviolet

Arman Cingöz^{1*}, Dylan C. Yost^{1*}, Thomas K. Allison¹, Axel Ruehl^{2†}, Martin E. Fermann², Ingmar Hartl² & Jun Ye¹

The development of the optical frequency comb (a spectrum consisting of a series of evenly spaced lines) has revolutionized metrology and precision spectroscopy owing to its ability to provide a precise and direct link between microwave and optical frequencies^{1,2}. A further advance in frequency comb technology is the generation of frequency combs in the extreme-ultraviolet spectral range by means of high-harmonic generation in a femtosecond enhancement cavity^{3,4}. Until now, combs produced by this method have lacked sufficient power for applications, a drawback that has also hampered efforts to observe phase coherence of the high-repetition-rate pulse train produced by high-harmonic generation, which is an extremely nonlinear process. Here we report the generation of extreme-ultraviolet frequency combs, reaching wavelengths of 40 nanometres, by coupling a high-power near-infrared frequency comb⁵ to a robust femtosecond enhancement cavity. These combs are powerful enough for us to observe single-photon spectroscopy signals for both an argon transition at 82 nanometres and a neon transition at 63 nanometres, thus confirming the combs' coherence in the extreme ultraviolet. The absolute frequency of the argon transition has been determined by direct frequency comb spectroscopy. The resolved ten-megahertz linewidth of the transition, which is limited by the temperature of the argon atoms, is unprecedented in this spectral region and places a stringent upper limit on the linewidth of individual comb teeth. Owing to the lack of continuous-wave lasers, extreme-ultraviolet frequency combs are at present the only promising route to extending ultrahigh-precision spectroscopy to the spectral region below 100 nanometres. At such

wavelengths there is a wide range of applications, including the spectroscopy of electronic transitions in molecules⁶, experimental tests of bound-state and many-body quantum electrodynamics in singly ionized helium and neutral helium^{7–9}, the development of next-generation 'nuclear' clocks^{10–12} and searches for variation of fundamental constants¹³ using the enhanced sensitivity of highly charged ions¹⁴.

Techniques developed to control a train of ultrashort pulses in the frequency domain have led to rapid advancements not only in ultrahigh-precision metrology¹ but also in the generation of attosecond pulses for time-resolved studies¹⁵. This close relationship between time and frequency techniques continues with the development of the extreme-ultraviolet (XUV) frequency combs, in which high-harmonic generation (HHG), a standard technique in attosecond physics, is used to produce phase-coherent XUV radiation. In conventional HHG, a single infrared pulse generates a burst of attosecond pulses separated by half cycles of the driving laser field, resulting in an odd-harmonic spectrum (Fig. 1). By contrast, in intracavity HHG, a phase-coherent infrared pulse train is used to produce a train of such bursts that repeat at the repetition frequency of the fundamental comb. This new temporal structure is responsible for the much finer frequency comb at each harmonic order. We anticipate that the XUV frequency comb will allow high-precision characterization of the HHG process and provide further insights into attosecond physics.

The existence of the XUV comb structure is critically dependent on the phase coherence of the HHG process. The temporal coherence of HHG in an isolated pulse has been studied extensively^{16,17}, but the

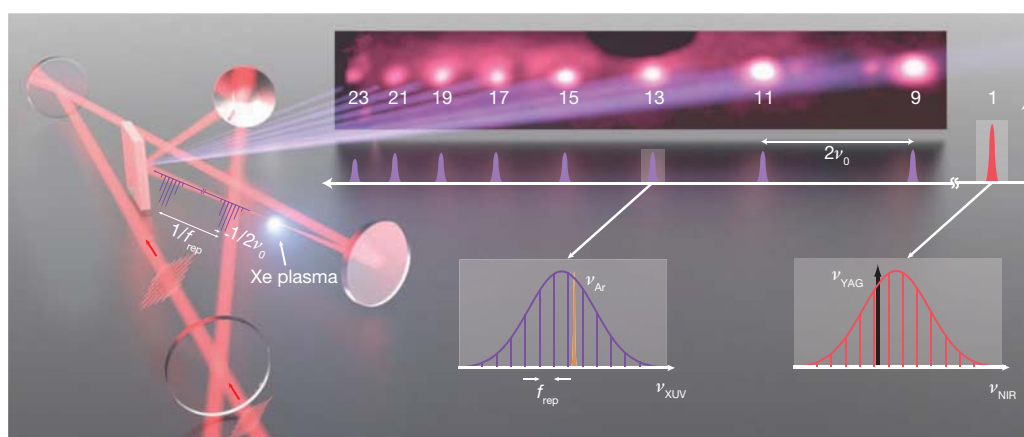


Figure 1 | Intracavity high-harmonic generation. An infrared frequency comb is passively amplified in a power build-up cavity with a mirror/diffraction grating hybrid output coupler for the XUV range²⁶. Xenon gas is introduced at the tight focus of the cavity. The resultant HHG spectrum (top inset) from a single infrared pulse consists of odd harmonics of the infrared carrier frequency, ν_0 . By using an infinite train of pulses, we create a new comb

structure in every harmonic at the repetition frequency (f_{rep}) of the pulse train (lower insets). This comb structure is stabilized using a continuous-wave laser (frequency, ν_{YAG}) and slowly scanned over an argon transition (frequency, ν_{Ar}) at 82 nm. ν_{XUV} and ν_{NIR} respectively represent frequencies in the XUV and near-infrared spectral regions.

¹JILA, National Institute of Standards and Technology and University of Colorado, Department of Physics, University of Colorado, Boulder, Colorado 80309-0440, USA. ²IMRA America Inc., 1044 Woodridge Avenue, Ann Arbor, Michigan 48105, USA. [†]Present address: Institute for Lasers, Life and Biophotonics, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081HV Amsterdam, The Netherlands.

*These authors contributed equally to this work.

requirements are more stringent for XUV comb generation, where the phase coherence must be maintained over many consecutive pulses. This issue has been a topic of investigation since the first demonstrations of intracavity HHG^{3,4}, in which only the phase coherence of the third harmonic was verified. In subsequent work, pulse-to-pulse temporal coherence of the seventh harmonic at 152 nm, below the ionization threshold, was demonstrated interferometrically¹⁸. The pulse-to-pulse phase coherence of above-threshold harmonics in the XUV range was first demonstrated in ref. 8. This coherence was critical for two-pulse Ramsey spectroscopy of helium at 51 nm with a systematic uncertainty of 6 MHz (ref. 8) in a widely tunable system¹⁹. We note that this approach differs fundamentally from the work presented here. The two-pulse Ramsey sequence is a phase measurement and, as a result, is susceptible to systematic phase errors between the two pulses. These errors lead to apparent frequency shifts. Thus, all sources of phase error must be considered and characterized before an absolute frequency measurement can be made. By contrast, a frequency comb consisting of an infinite pulse train is an optical frequency synthesizer with equally spaced frequency comb teeth, defined by only two numbers: the repetition frequency (f_{rep}) of the pulse train and the carrier envelope offset frequency (f_{ceo}). Thus, the phase instabilities can lead to broadening of the comb teeth, but the absolute frequencies of the XUV comb remain simply and robustly determined. It is precisely this characteristic that led to the recent frequency comb revolution in metrology even though multipulse Ramsey spectroscopy in the visible and infrared spectral regions has existed since the late 1970s²⁰. In this Letter, we demonstrate the extension of frequency comb spectroscopy to the XUV spectral region, which will lead to many applications in short-wavelength metrology.

In intracavity HHG, a high-average-power infrared frequency comb⁵ is coupled to a dispersion-controlled optical cavity (Fig. 1), which allows for passive coherent pulse build-up at the full repetition rate of 154 MHz (Supplementary Information). The high repetition rate of intracavity HHG should in principle provide increased average flux relative to conventional single-pass HHG systems, which typically use lasers with repetition frequencies and average powers that are smaller by orders of magnitude. However, until now the presence of the optical cavity has presented challenges to matching the average flux of these systems, owing to optical damage problems associated with the large average intracavity power and to the nonlinear-response of the HHG medium. At the high intensities and gas densities required to optimize the XUV flux, the plasma density at the intracavity focus reaches levels that can lead to various nonlinearities such as optical bistability, self-phase modulation and pulse distortion. Recently, these effects have been studied in detail with both simulations and experiments^{21,22}. These effects, amplified by the narrow cavity linewidth, clamp the achievable intracavity peak power and lead to instabilities in the lock between the comb and the cavity, limiting the attainable XUV flux and coherence time. Thus, a robust means of mitigating these effects is to decrease the finesse. However, this requires that a more powerful laser be used to excite the enhancement cavity to keep the intracavity pulse energy high. This obstacle has been overcome by recent advances in high-repetition-rate, high-average-power chirped-pulse amplified Yb:fibre lasers^{5,23–25}.

Our ability to reach high intracavity pulse energies while minimizing the deleterious effects of a large plasma density has resulted in the generation of unprecedented power in high harmonics spectrally separated for spectroscopy. We measure the power in the fifteenth harmonic to be 21 μW with xenon as the target gas (Fig. 2), which corresponds to more than 200 μW per harmonic produced in the cavity. This is an increase of more than an order of magnitude over previous intracavity HHG results^{26,27}. Similar outcoupled power (not spectrally separated) was recently reported²⁸. We have also observed shorter-wavelength radiation by using krypton as the target gas for HHG. Owing to the higher ionization potential of krypton, we were able to detect the twenty-fifth and twenty-seventh harmonic orders

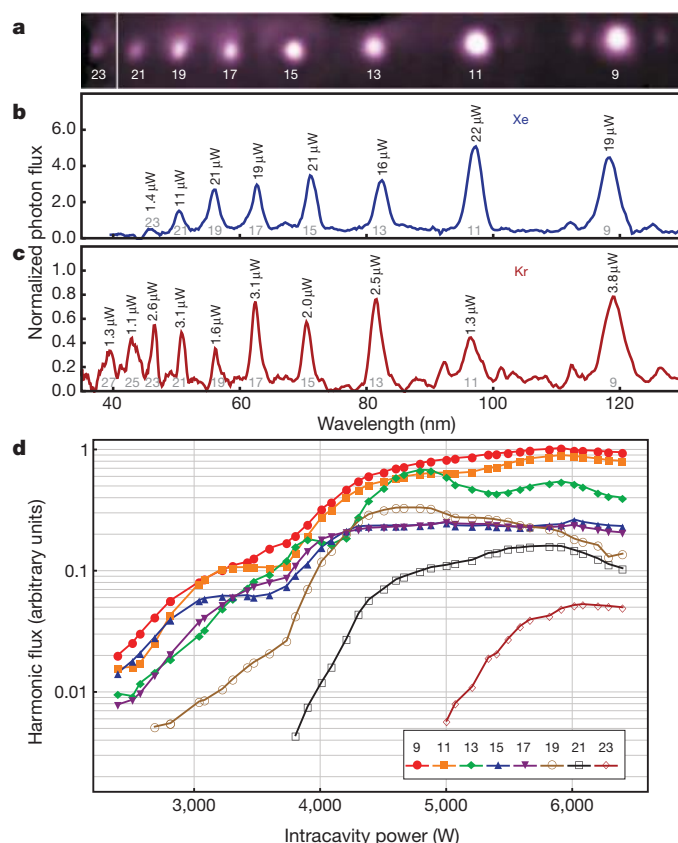


Figure 2 | Power scaling results. **a**, Characteristic image of the harmonics in xenon dispersed on a plate coated with fluorescent material. The threshold of the image was adjusted for the twenty-third harmonic to make it more visible. **b**, **c**, Vertically integrated intensity from a charge-coupled-device (CCD) image for xenon (**b**) and krypton (**c**) target gas, and corresponding integrated harmonic powers (Methods Summary). The observed widths of the harmonic orders are given by the instrument wavelength resolution. The photon flux has been normalized by the spectral bandwidth. **d**, Harmonic powers generated in xenon as functions of infrared intracavity power, showing saturation at powers greater than 5.5 kW. The oscillations visible for some harmonics are due to quantum path interference¹⁸.

with 8 kW of intracavity power (peak intensity of $9 \times 10^{13} \text{ W cm}^{-2}$) (Fig. 2c), extending intracavity HHG to photon energies above 30 eV. As shown in Fig. 2d, there are only modest improvements in power for most harmonics as the intracavity infrared power is increased above 5.5 kW. As the plasma density increases, depletion of the on-axis neutral atomic density and a decrease in phase matching cause the XUV power clamp. These results show that future gains in harmonic power in a similar system will most probably be achieved by increasing the focal area at the interaction region.

To demonstrate the comb structure of XUV radiation, we have spectrally resolved the $3s^23p^6 J=0 \rightarrow 3s^23p^55d J=1$ electric dipole (E1) transition in argon with an upper-state energy of $121,932.8 \text{ cm}^{-1}$ (corresponding to a wavelength of $\sim 82 \text{ nm}$) and the $2s^22p^6 J=0 \rightarrow 2s^22p^54s J=1$ E1 transition in neon with an upper-state energy of $159,534.6 \text{ cm}^{-1}$ ($\sim 63 \text{ nm}$) by means of resonance fluorescence spectroscopy. These transitions lie within the thirteenth and seventeenth harmonic bandwidths of our source, respectively. To reduce the Doppler width of the transition below the repetition frequency of the comb, we use pulsed supersonic atomic beams propagating orthogonally to the XUV beam. The average harmonic power available for spectroscopy is $\sim 1 \mu\text{W}$ owing to a single reflection from a mirror to steer the XUV beam into the interaction region and several pinholes to isolate the central part of the beam that contains the phase-coherent short trajectories¹⁸. Moreover, owing to the electric dipole nature of these transitions, only a single comb tooth with $\sim 10^{-5}$ of the total power,

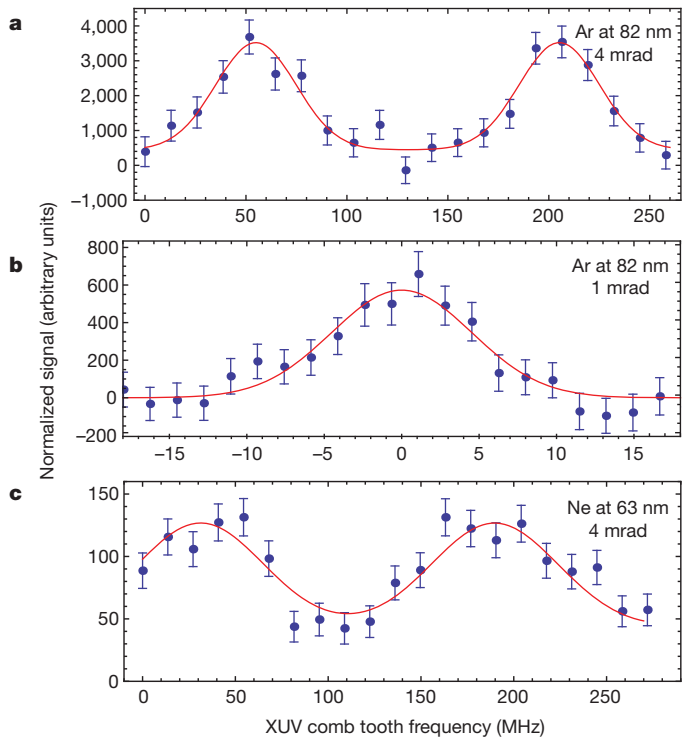


Figure 3 | Atomic fluorescence signal. **a, b,** Photomultiplier signals as functions of an XUV comb tooth frequency for argon beam collimation angles of ~ 4 mrad (**a**) and ~ 1 mrad (**b**). The signal is normalized to the incident XUV power and the scattered photon background is subtracted (Methods Summary). Error bars (1 s.d.) are determined from the shot noise of counted fluorescence photons. The red curves are least-squares fits to a Gaussian model. **c,** Signal for the neon transition with a collimation angle of ~ 4 mrad. The contrast and signal-to-noise ratio are lower owing to the increased Doppler width and the smaller absorption cross-section, respectively.

10 pW, contributes to the signal. Nevertheless, we have achieved a ratio of signal to scattered photon background in the range of 0.5–1 using a photomultiplier tube to detect directly the XUV spontaneous emission to the ground state.

The fundamental infrared comb is controlled by phase-locking a comb tooth to a continuous-wave neodymium-doped yttrium

aluminium garnet (Nd:YAG) laser by feedback onto the laser cavity length². The Nd:YAG laser frequency (ν_{YAG}) is referenced to a molecular iodine transition and serves as an absolute optical frequency marker with an uncertainty of 5 kHz (ref. 29). The comb offset frequency, f_{ceo} , has a free-running linewidth of 15 kHz (ref. 5) and is adjusted to optimize intracavity power, but is otherwise unlocked. Its value is deduced from the knowledge of ν_{YAG} and f_{rep} , which we measure against the hydrogen maser at the US National Institute of Standards and Technology (Supplementary Information). Figure 3 shows the fluorescence signal as the phase-locked-loop offset frequency, δ , is slowly changed and XUV comb teeth are scanned across the atomic resonance for two different collimation angles of the atomic beam. With the scattered photon background subtracted, the signal contrast is 100%, confirming the sharp comb structure. Least-squares fits for the argon data reveal respective Gaussian full-widths at half-maximum of 47 ± 5 and 11 ± 1 MHz for the two angles. These results are consistent with what would be expected from the geometric collimation. Thus, the observed linewidth places only an upper bound on the comb teeth linewidth, and this bound is less than 10 MHz. The statistical uncertainty in the determination of the line centre for the narrower line shape is 500 kHz, demonstrating an unprecedented fractional frequency precision of better than 2×10^{-10} in the XUV range.

To determine the absolute frequency of the argon transition, it is necessary to determine the comb tooth number in addition to f_{rep} and f_{ceo} . This requires making several measurements at different values of f_{rep} and keeping track of the change in the comb tooth number (Supplementary Information). Figure 4 shows the results of ten measurements at different values of f_{rep} chosen to provide absolute comb mode determination. The data analysis is carried out by noting that the fixed point of the infrared comb (comb tooth phase locked to the Nd:YAG laser) is multiplied by the HHG process to create corresponding fixed points for every harmonic comb. These fixed points are offset from $q\nu_{\text{YAG}}$ by $\Delta = q\delta(\text{modulo } f_{\text{rep}})$, where q is the harmonic order. Thus, the argon transition frequency is

$$\nu_{\text{Ar}} = n f_{\text{rep}} + 13 f_{\text{ceo}} = m f_{\text{rep}} + 13 \nu_{\text{YAG}} + \Delta$$

where n is the absolute comb tooth number and m is the difference in comb tooth number between the thirteenth harmonic fixed point and the actual comb tooth probing the argon transition. This point of view not only makes the analysis independent of drifts in f_{ceo} , but also highlights the direct frequency link between the infrared and the higher-harmonic combs.

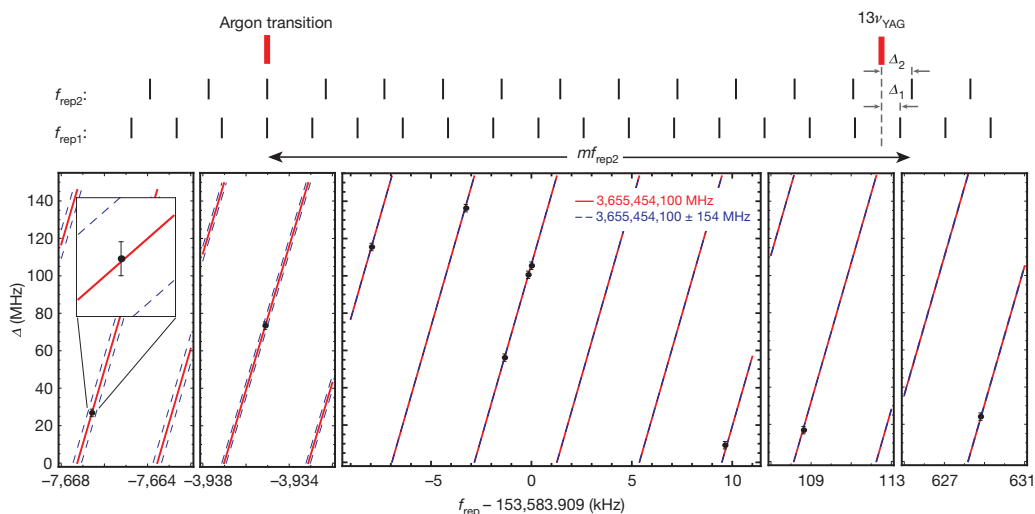


Figure 4 | Absolute frequency determination. Measurement of the resonance centre frequency at ten different values of f_{rep} parameterized by the offset, Δ , of the thirteenth harmonic fixed point from $13\nu_{\text{YAG}}$. The top part of the figure schematically illustrates two different measurements at $f_{\text{rep}1}$ and $f_{\text{rep}2}$, with corresponding respective offsets of Δ_1 and Δ_2 . The short black lines

represent the locations of individual frequency comb teeth. In the main plot, the solid red lines show the expected locations of the resonance centres for the extracted comb tooth number, m , and the pairs of blue dashed lines show the expected locations for $m \pm 1$. Inset, magnified version of a measurement, demonstrating that the comb tooth number is unambiguously determined.

Once m is determined, we correct the final value for residual Doppler shifts arising from imperfect orthogonal alignment of the laser and the atomic beam. This is done by comparing the measured transition frequency of a pure argon beam (velocity of $520 \pm 20 \text{ m s}^{-1}$) with that of a helium–argon mixture (5:1 ratio of partial pressures; velocity of $740 \pm 60 \text{ m s}^{-1}$) and translating the collimation slit until the two measured frequencies match, ensuring orthogonality between the laser and the atomic beams. Owing to the small fractional change in the velocity of the two beams, the uncertainty in the procedure is $\sim 3 \text{ MHz}$. All other systematic shifts such as recoil, density and Stark shifts, as well as the contribution due to the uncertainty in v_{YAG} , are estimated to be smaller than 1 MHz . We have also binned our data for different laser operation conditions and find no systematic shifts at the 2-MHz level for a 30% change in xenon backing pressure or a 20% change in the intracavity power for HHG, again highlighting the robustness of the XUV comb for frequency measurements. The final absolute frequency for this transition is $3,655,454,073 \pm 3 \text{ MHz}$, which agrees well with previous measurements that have an uncertainty of 2.3 GHz (ref. 30).

Now that the XUV frequency comb has been developed into a robust tool, it is imperative to determine its ultimate coherence. The current upper bound on the thirteenth harmonic comb linewidth is less than 10 MHz , limited by the atomic thermal motion. We plan to implement a heterodyne beat measurement between two intracavity HHG sources to overcome this limitation, improving the phase noise characterization by orders of magnitude. With these developments, ultrahigh-precision spectroscopy in the XUV range is firmly within our grasp, enabling a wide range of applications.

METHODS SUMMARY

The Yb:fibre comb delivers 120-fs pulses at a 154-MHz repetition frequency with a maximum average power of 80 W (ref. 5). For the studies presented here, we run the laser at $\sim 30 \text{ W}$. Target gas is injected at the intracavity focus using a glass nozzle with a $100\text{-}\mu\text{m}$ aperture and a backing pressure of $\sim 2 \text{ atm}$. The XUV light is coupled out of the cavity using an intracavity diffraction grating with 10% efficiency²⁶.

The power of the thirteenth harmonic is measured with an XUV photodiode. The power in the other harmonics is determined using a plate coated with sodium salicylate, which fluoresces at 420 nm when excited by XUV radiation, with uniform quantum efficiency between 40 and 100 nm (ref. 18). The fluorescence is imaged using a CCD camera and calibrated using the thirteenth harmonic, which can be measured with either the photodiode or the CCD. The estimated calibration uncertainty is $20\text{--}30\%$.

The supersonic argon source consists of a pulsed valve nozzle (0.5-mm diameter), a 0.5-mm -diameter skimmer and a secondary slit located 33 cm from the skimmer. Both the slit width and slit position can be adjusted to study the Doppler width and shifts. The valve generates $500\text{-}\mu\text{s}$ -long pulses at a repetition rate of 20 Hz with a backing pressure of 0.5 atm .

The fluorescence is detected with a photomultiplier tube and pulse counter. The counter has two gates: one coincident with the argon pulse arrival time, to measure the signal; and another located between argon pulses, to measure the background. Most of the laser light passes through the interaction region unimpeded and is detected with a second XUV photodiode, which provides the normalization signal. Individual frequency scans take $\sim 150 \text{ s}$. Results presented in Fig. 3 are averages of $3\text{--}10$ scans with an acquisition time of $8\text{--}25 \text{ min}$.

Received 9 September; accepted 3 November 2011.

1. Udem, T., Holzwarth, R. & Hänsch, T. W. Optical frequency metrology. *Nature* **416**, 233–237 (2002).
2. Cundiff, S. T. & Ye, J. Femtosecond optical frequency combs. *Rev. Mod. Phys.* **75**, 325–342 (2003).
3. Jones, R. J., Moll, K. D., Thorpe, M. J. & Ye, J. Phase-coherent frequency combs in the vacuum ultraviolet via high-harmonic generation inside a femtosecond enhancement cavity. *Phys. Rev. Lett.* **94**, 193201 (2005).

4. Gohle, C. *et al.* A frequency comb in the extreme ultraviolet. *Nature* **436**, 234–237 (2005).
5. Ruehl, A., Marcinkevicius, A., Fermann, M. E. & Hartl, I. 80 W , 120 fs Yb-fiber frequency comb. *Opt. Lett.* **35**, 3015–3017 (2010).
6. Merk, F. & Softley, T. P. Final-state interactions in the zero-kinetic-energy-photoelectron spectrum of H_2 . *J. Chem. Phys.* **96**, 4149–4156 (1992).
7. Herrmann, M. *et al.* Feasibility of coherent XUV spectroscopy on the $1\text{S}\text{--}2\text{S}$ transition in singly ionized helium. *Phys. Rev. A* **79**, 052505 (2009).
8. Kandula, D. Z., Gohle, C., Pinkert, T. J., Ubachs, W. & Eikema, K. S. E. Extreme ultraviolet frequency comb metrology. *Phys. Rev. Lett.* **105**, 063001 (2010).
9. Eyler, E. E. *et al.* Prospects for precision measurements of atomic helium using direct frequency comb spectroscopy. *Eur. Phys. J. D* **48**, 43–55 (2008).
10. Peik, E. & Tamm, C. Nuclear laser spectroscopy of the 3.5 eV transition in Th-229. *Europhys. Lett.* **61**, 181–186 (2003).
11. Rellergert, W. G. *et al.* Constraining the evolution of the fundamental constants with a solid-state optical frequency reference based on the ^{229}Th nucleus. *Phys. Rev. Lett.* **104**, 200802 (2010).
12. Campbell, C. J., Radnaev, A. G. & Kuzmich, A. Wigner crystals of ^{229}Th for optical excitation of the nuclear isomer. *Phys. Rev. Lett.* **106**, 223001 (2011).
13. Murphy, M. T., Webb, J. K. & Flambaum, V. V. Further evidence for a variable fine-structure constant from Keck/HIRES QSO absorption spectra. *Mon. Not. R. Astron. Soc.* **345**, 609–638 (2003).
14. Berengut, J. C., Dzuba, V. A., Flambaum, V. V. & Ong, A. Electron-hole transitions in multiply charged ions for precision laser spectroscopy and searching for variations in α . *Phys. Rev. Lett.* **106**, 210802 (2011).
15. Krausz, F. & Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **81**, 163–234 (2009).
16. Bellini, M. *et al.* Temporal coherence of ultrashort high-order harmonic pulses. *Phys. Rev. Lett.* **81**, 297–300 (1998).
17. Mairesse, Y. *et al.* Attosecond synchronization of high-harmonic soft X-rays. *Science* **302**, 1540–1543 (2003).
18. Yost, D. C. *et al.* Vacuum-ultraviolet frequency combs from below-threshold harmonics. *Nature Phys.* **5**, 815–820 (2009).
19. Pinkert, T. J. *et al.* Widely tunable extreme UV frequency comb generation. *Opt. Lett.* **36**, 2026–2028 (2011).
20. Eckstein, J. N., Ferguson, A. I. & Hänsch, T. W. High-resolution two-photon spectroscopy with picosecond light pulses. *Phys. Rev. Lett.* **40**, 847–850 (1978).
21. Allison, T. K., Cingöz, A., Yost, D. C. & Ye, J. Extreme nonlinear optics in a femtosecond enhancement cavity. *Phys. Rev. Lett.* **107**, 183903 (2011).
22. Carlson, D. R., Lee, J., Mongelli, J., Wright, E. M. & Jones, R. J. Intracavity ionization and pulse formation in femtosecond enhancement cavities. *Opt. Lett.* **36**, 2991–2993 (2011).
23. Hartl, I. *et al.* Cavity-enhanced similariton Yb-fiber laser frequency comb: $3 \times 10^{14} \text{ W/cm}^2$ peak intensity at 136 MHz . *Opt. Lett.* **32**, 2870–2872 (2007).
24. Schibli, T. R. *et al.* Optical frequency comb with submillihertz linewidth and more than 10 W average power. *Nature Photon.* **2**, 355–359 (2008).
25. Eidam, T. *et al.* Femtosecond fiber CPA system emitting 830 W average output power. *Opt. Lett.* **35**, 94–96 (2010).
26. Yost, D. C., Schibli, T. R. & Ye, J. Efficient output coupling of intracavity high-harmonic generation. *Opt. Lett.* **33**, 1099–1101 (2008).
27. Ozawa, A. *et al.* High harmonic frequency combs for high resolution spectroscopy. *Phys. Rev. Lett.* **100**, 253901 (2008).
28. Lee, J., Carlson, D. & Jones, R. J. Optimizing intracavity high harmonic generation for XUV fs frequency combs. *Opt. Express* **19**, 23315–23326 (2011).
29. Ye, J., Ma, L.-S. & Hall, J. L. Molecular iodine clock. *Phys. Rev. Lett.* **87**, 270801 (2001).
30. Minnhagen, L. Spectrum and the energy levels of neutral argon, Ar I. *J. Opt. Soc. Am.* **63**, 1185–1198 (1973).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. L. Hall for the use of an iodine-stabilized laser, M. D. Swallows for the assistance with the hydrogen maser frequency transfer, and S. T. Cundiff and A. Foltynowicz for reading a draft of the manuscript. This research is funded by the DARPA, AFOSR, NIST and NSF. A.C. and T.K.A. are National Research Council postdoctoral fellows. A.R. acknowledges funding from the Alexander von Humboldt Foundation (Germany).

Author Contributions A.C., D.C.Y., T.K.A. and J.Y. conceived of, designed and carried out the XUV power and spectroscopy measurements. A.R., M.E.F. and I.H. designed and built the Yb-fibre laser. All authors discussed the results and worked on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.C. (acingo@jila.colorado.edu) or J.Y. (junye@jila.colorado.edu).

Nonlinear material behaviour of spider silk yields robust webs

Steven W. Cranford^{1,2}, Anna Tarakanova^{1,2,3}, Nicola M. Pugno⁴ & Markus J. Buehler^{1,2,5}

Natural materials are renowned for exquisite designs that optimize function, as illustrated by the elasticity of blood vessels, the toughness of bone and the protection offered by nacre^{1–5}. Particularly intriguing are spider silks, with studies having explored properties ranging from their protein sequence⁶ to the geometry of a web⁷. This material system⁸, highly adapted to meet a spider's many needs, has superior mechanical properties^{9–15}. In spite of much research into the molecular design underpinning the outstanding performance of silk fibres^{1,6,10,13,16,17}, and into the mechanical characteristics of web-like structures^{18–21}, it remains unknown how the mechanical characteristics of spider silk contribute to the integrity and performance of a spider web. Here we report web deformation experiments and simulations that identify the nonlinear response of silk threads to stress—involving softening at a yield point and substantial stiffening at large strain until failure—as being crucial to localize load-induced deformation and resulting in mechanically robust spider webs. Control simulations confirmed that a nonlinear stress response results in superior resistance to structural defects in the web compared to linear elastic or elastic–plastic (softening) material behaviour. We also show that under distributed loads, such as those exerted by wind, the stiff behaviour of silk under small deformation, before the yield point, is essential in maintaining the web's structural integrity. The superior performance of silk in webs is therefore not due merely to its exceptional ultimate strength and strain, but arises from the nonlinear response of silk threads to strain and their geometrical arrangement in a web.

Although spider silk is used by spiders for many purposes, from wrapping prey to lining retreats^{22,23}, here we focus on silk's structural role in aerial webs and on how silk's material properties relate to web function. The mechanical behaviour of silk, like that of other biological materials, is determined by the nature of its constituent molecules and their hierarchical assembly into fibres^{13,16,17,24–26} (Supplementary Fig. 1). Spider webs themselves are characterized by a highly organized geometry that optimizes their function^{7,8,18–20}. To explore the contribution of the material characteristics to web function, we developed a web model with spiral and radial threads based on the geometry commonly found in orb webs¹. The silk material behaviour was parameterized from atomistic simulations of dragline silk from the species *Nephila clavipes* (model A)^{16,17} (Fig. 1a, b) and validated against experiments¹⁰ (Methods Summary). Properties of silk can vary across evolutionary lineages by over 100% (refs 9, 27 and 28; Supplementary Information section 1), so we avoided species-specific silk properties and instead used a representative model to reflect the characteristic nonlinear stress–strain (σ – ϵ) behaviour of silk found in a web. The mechanical performance of individual silk threads has been previously investigated^{10,12,13}, and is in agreement with our model in terms of tensile deformation behaviour.

It is rare to see a perfectly intact web—debris, attack or unstable anchorage lead to loss of threads (see inset to Fig. 1c)—but the structure usually remains functional for a spider's use. We assessed a web's ability to tolerate defects by removing web sections (silk threads) and applying a local load (Fig. 1c). Removal of up to 10% of threads, at different locations relative to the load, had little impact on the web's response; in fact, the ultimate load capacity increased by 3–10% with the introduction of defects (Fig. 1c). We observed in all cases that failure is limited to the thread to which the force is applied. Loading of a spiral thread resulted in relatively isolated web distortion (Fig. 1e), whereas loading of a radial thread (Fig. 1f) resulted in larger deformation (about 20% more deflection and about 190% increase in energy dissipation; Fig. 1d). But in both cases, failure was localized (Fig. 1e, f). A comparative study of loading radial versus spiral threads demonstrated that the web's structural performance is dominated by the properties of the stiffer and stronger radial dragline silk (with the force required to break radial threads within the web approximately 150% higher), suggesting that the spiral threads play non-structural roles (such as capturing prey).

In situ experiments on a garden spider (*Araneus diadematus*) web (Fig. 1e, f) were in qualitative agreement with the simulations: they confirmed the prediction that failure is localized when loading either a spiral or a radial thread. Complementing these findings, we used our atomistic silk model^{16,17} to connect the stress states in the web (Fig. 2a, top row) with molecular deformation mechanisms in the threads (Fig. 1a). Under loading and immediately before failure, most radial threads in the structure exhibited deformation states equivalent to the yield regime (regime II in Fig. 1a), where the presence of polymer-like semi-amorphous regions permits entropic unfolding of the silk nanocomposite under relatively low stress^{16,17,29}. Once unfolding is complete, the system stiffens as stress is transferred to relatively rigid β -sheet nanocrystals¹⁷ (regimes III–IV in Fig. 1a); it finally fails, at the thread where force is applied, because the applied stress is sufficient to rupture the nanocrystals.

Simulation and experiment both indicated that localized failure is a universal characteristic of spider webs. It is unresolved whether this behaviour is unique to silk-like materials or a result of the web's architecture (that is, a property of the construction material or of the structural design). We therefore systematically compared the response of webs constructed from three different types of fibres with distinct mechanical behaviour (Fig. 2a, left panels): in addition to fibres with the atomistically derived stress–strain behaviour of dragline silk (model A), we used idealized engineered fibres that exhibited either linear elastic behaviour (model A') or elastic–perfectly plastic behaviour that involves severe softening (plastic yield) (model A''). In all cases, we loaded one of the radial threads and assumed that the failure stress (about 1,400 MPa) and strain (about 67%) of silk threads are constant, so that any changes in deformation behaviour (Fig. 2a, right

¹Laboratory for Atomistic and Molecular Mechanics (LAMM), Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. ²Center for Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. ³Department of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA. ⁴Laboratory of Bio-Inspired Nanomechanics “Giuseppe Maria Pugno”, Department of Structural Engineering and Geotechnics, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy. ⁵Center for Computational Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

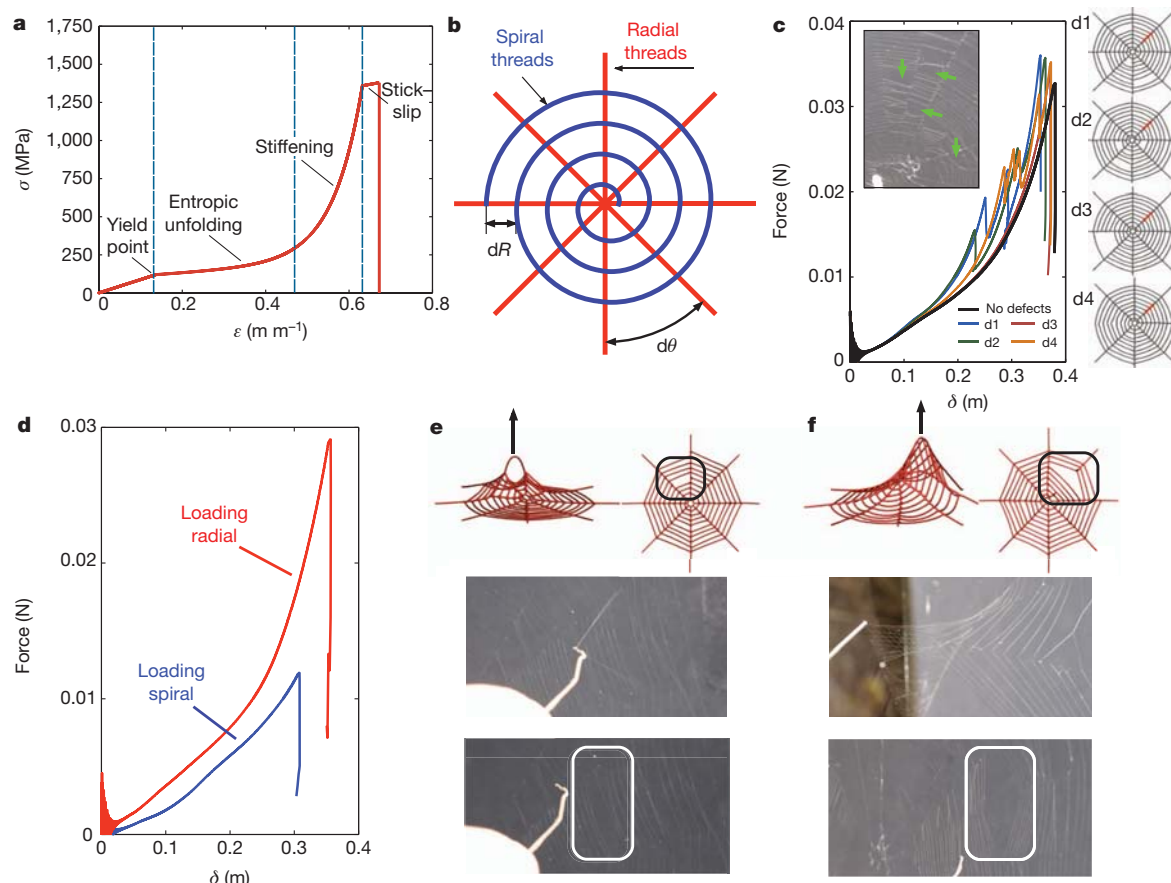


Figure 1 | Material behaviour of dragline spider silk, web model, and behaviour of webs under load. **a**, Derived stress–strain (σ – ϵ) behaviour of dragline silk, parameterized from atomistic simulations and validated against experiments^{16,17}. There are four distinct regimes characteristic of silk^{16,17}. I, stiff initial response governed by homogeneous stretching; II, entropic unfolding of semi-amorphous protein domains; III, stiffening regime as molecules align and load is transferred to the β -sheet crystals; and IV, stick-slip deformation of β -sheet crystals¹⁶ until failure. **b**, Schematic of web model, approximated by a continuous spiral (defined by dR) supported by eight regular radial silk threads (defined by $d\theta$), typical of orb webs⁷. **c**, Force–displacement curves for loading a

defective web (results for model A; loaded region shown in red). Case studies include missing spiral segments (d1 to d3) and a missing radial thread (d4). The inset to **c** shows the *in situ* orb web as discovered, containing many defects (marked by green arrows). **d**, Force–displacement behaviour of web, comparing the loading of a single radial thread and a single spiral thread (model A). **e**, Loading of a spiral thread results in small web deformation. **f**, Loading applied at radial threads results in an increase in web deformation. In both cases (**e** and **f**) failure is isolated to the pulled thread in simulation and experiment, restricting damage to a small section of the web (indicated by white rectangles).

panels) and web damage (Fig. 2a, middle panels) would be a direct result of differences in the stress–strain behaviour of the fibres. In the case of a web comprised of natural dragline silk (top panels of Fig. 2a), all radial threads contributed partially to the resistance to loading, but the fact that the material suddenly softened at the yield point, which immediately reduced the initial modulus (about 1,000 MPa) by around 80%, ensured that only the loaded radial thread entered regime III and began to stiffen before it finally failed. With linear elastic material behaviour (middle panels of Fig. 2a), the loaded radial thread was still subjected to the bulk of the load; but adjacent radial threads bore a higher fraction of the ultimate load, which resulted in a greater delocalization of damage upon failure. With elastic–perfectly plastic behaviour (bottom panels of Fig. 2a), the softening of radial threads enhanced the load distribution even more throughout the web and thereby greatly increased the damage zone once failure occurred. The increased contribution of the auxiliary radial threads to load resistance as we moved from the natural to linear elastic to elastic–perfectly plastic behaviour resulted in 34% higher maximum strength, but 30% less displacement at failure (Fig. 2b).

The above simulations using atomistically derived silk properties (model A) assume that the spiral threads and radial threads are made of dragline silk and behave identically, except for differences arising from their different thread diameters. But in real spider webs, spiral threads are composed of more compliant and extensible viscid silk (for

example, a failure strain of around 270% for the species *Araneus diadematus*¹). To explore the effect of different silks making up the spiral and radial threads, we introduced empirically parameterized viscid spiral threads¹ (model B) and found that the results were only marginally affected (Fig. 2b). We also used a model in which we parameterized both spiral and radial threads according to empirical data¹ (model C), subjected this model to the same loading conditions and systematically compared its performance against that of models with linear elastic (model C') and elastic–perfectly plastic behaviours (model C''). We found similar web responses and although the web made from natural silk is weaker, it still localizes damage near the loaded region (Supplementary Information section 5).

To explore global loading responses, we subjected the web models to a homogeneously distributed wind load with effective wind speeds up to 70 m s^{-1} (a threshold at which all models fail). The system-level deflection curves (Fig. 2c) reflect the mechanical behaviour of the radial threads, which ultimately transferred load to the web's anchoring points. Although the spiral threads underwent increased deflection and captured more of the wind load owing to their larger exposed length, they were effectively pinned to the much stiffer dragline radial threads that limit web deflection (Fig. 2c, Supplementary Information section 8). For wind speeds less than 10 m s^{-1} there was little difference between the models (Fig. 2c, Supplementary Information section 8) and deflections are $<12\%$ of the total span of the web. We attributed

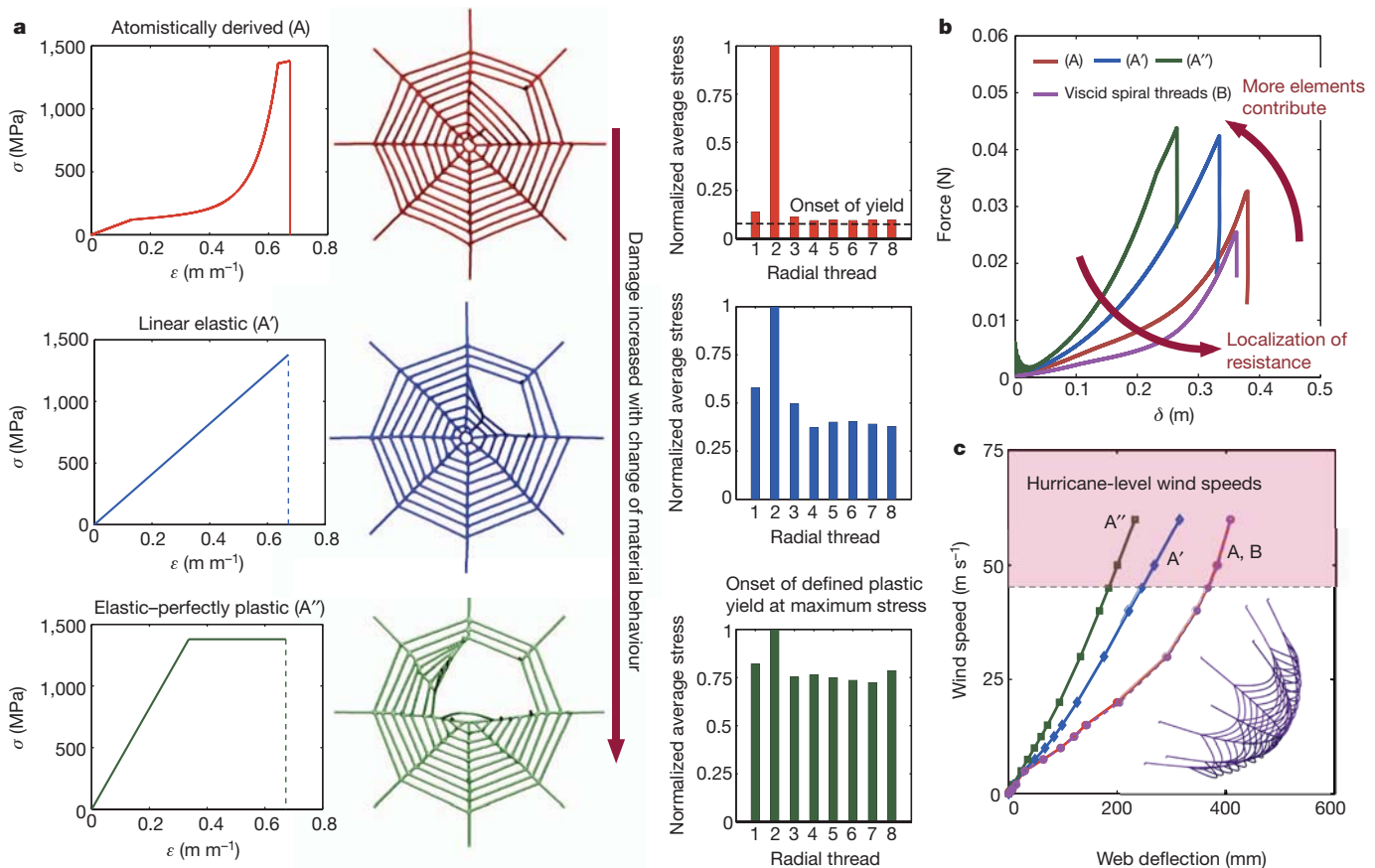


Figure 2 | Web response for varied silk behaviour under targeted (local) and distributed (global) loading. **a**, Comparison of failure for derived dragline silk, linear elastic and elastic-perfectly plastic behaviours (left, models A, A' and A''). Comparison of failure (centre) confirms localized stresses and minimized damage for the natural nonlinear stiffening silk behaviour. The average stress of each radial thread (bar plots, right) reflects the nonlinear deformation states in the silk. When load is applied locally to a radial thread, other radial threads not subject to applied force reach a stress corresponding to the onset of yielding

(that is, regime II in Fig. 1a). The elastic-perfectly plastic behaviour leads to an almost homogeneous distribution of stress. **b**, Force-displacement curves for varying material behaviours (models A, A' and A'' and model B). **c**, Web behaviour under distributed (global) wind loading. The plot shows a comparison of the wind-deflection behaviour (models A, A', A'' and B). The initial high stiffness of natural dragline silk enhances the structural integrity of the web under such loading. Failure of all webs occurs at wind speeds in excess of 60 m s^{-1} .

this relatively uniform structural rigidity of the web to the initial stiffness of the dragline silk before yield (Fig. 1a). Under higher wind loads, the softening behaviour of dragline silk at moderate deformation resulted in significant web deflection that was greater than the deflections seen with linear elastic and elastic-perfectly plastic material behaviour (Fig. 2c). We found that yield in the threads occurred at wind speeds exceeding around 5 m s^{-1} , defining a reasonable wind speed regime in which webs are operational.

Although all web models performed similarly under moderate global (wind) loading (Fig. 2c), the linear elastic and elastic-perfectly plastic models responded to targeted force application with a more catastrophic, brittle-like failure that resulted in significantly increased damage. Defining web damage as percentage of failed (broken) threads, we found that the damage of 2.5% for the natural silk behaviour increases sixfold to 15% for the elastic-perfectly plastic model (Fig. 2a, centre panel). Web performance under local loading was generalized by invoking quantized fracture mechanics³⁰, a theory that describes the failure mechanisms of discrete structures (such as a spider web) and adapted here to incorporate the material behaviours (Supplementary Information section 10). A generalized stress-strain behaviour, where $\sigma \propto \epsilon^\kappa$ (κ is a parameter that defines the nonlinear nature of the stress-strain relationship) treated with quantized fracture mechanics revealed that the size of the damaged zone in the proximity to a defect increases for materials that feature a softening behaviour (elastic-perfectly plastic behaviour), whereas a stiffening material (natural silk) results in a decrease of the damage zone (Fig. 3). This

is captured by a scaling law $\Omega(\alpha) = 1 - S^{2\alpha}$, defining the structural robustness Ω as the undamaged fraction of the web after failure. Here $\alpha = \kappa/(\kappa + 1)$ reflects the stress-strain response (linear elastic case when $\alpha = 1/2$; stiffening when α tends to 1 and softening when α tends to 0), and S is a system-dependent constant (independent of stress-strain relation). Our simulation results agreed with the predictions of quantized fracture mechanics (Supplementary Table 5) and confirmed that the relative size of the damage zone is a function of the material stress-strain relation and enhanced by the discreteness of the web (Supplementary Information section 10). This phenomenon is exemplified in spider webs (Figs 1 and 2), where the nonlinear stiffening behaviour (as α tends to 1) is essential for localizing damage and ensuring that a loaded thread becomes a sacrificial element while the majority of the web remains intact. Given the presumed metabolic effort required by the spider for rebuilding an entire web, localized failure is preferential as it does not compromise the structural integrity of the web (see Fig. 1c) and hence allows it to continue to function for prey capture in spite of the damage.

The remarkable strength, toughness and extensibility of individual spider silk threads are thus not the dominating properties that underpin the excellent structural performance of a spider web. Rather, it is the distinct nonlinear softening and subsequent stiffening of dragline silk that is essential to function, as it results in localization of damage to sacrificial threads in a web subjected to targeted (local) loading while minimizing web deformations under moderate wind (global) loading. Each regime of the nonlinear material behaviour of silk (Fig. 1a) thus

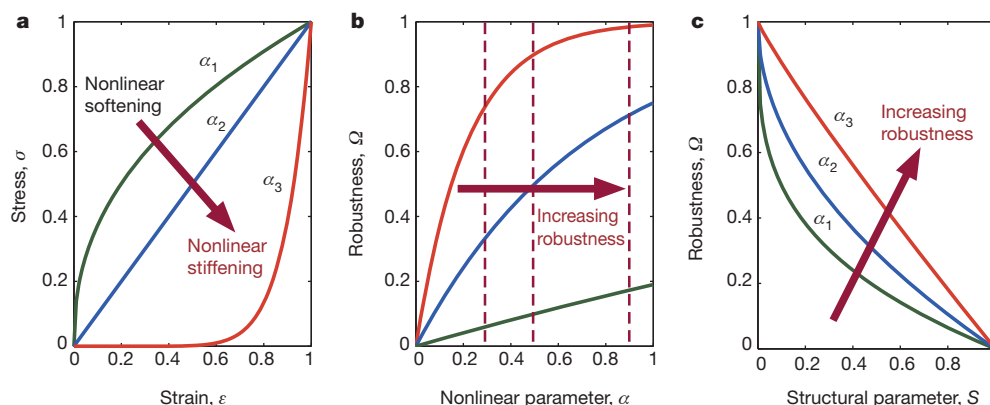


Figure 3 | Effects of stress–strain behaviour on structural robustness via quantized fracture mechanics. **a**, Plots of stress–strain curves (material behaviour) demonstrating the transition from softening to stiffening behaviours by the nonlinear parameter α ($\alpha_1 = 0.3$, $\alpha_2 = 0.5$, $\alpha_3 = 0.9$). **b**, Structural robustness Ω , defined as the undamaged fraction of the structure, versus α (dashed lines indicating α_1 , α_2 , α_3), given for three values of the system-

dependent constant S ($S_1 = 0.1$, $S_2 = 0.5$, $S_3 = 0.9$). S captures a range of material properties (such as fracture toughness), system geometry (that is, crack width or element length), and applied loading conditions. **c**, Ω versus S . Universally, the robustness increases with an increase in α ($\alpha_1 < \alpha_2 < \alpha_3$), implying that larger nonlinear stiffening results in larger structural robustness, and hence less damage.

plays a key part in defining the overall system response in a variety of environmental settings. Other natural silk threads used to form solid materials such as cocoons, rather than aerial webs, typically display different mechanical responses¹¹. Indeed, cocoon silk conforms closely to elastic–perfectly plastic behaviour, which is not suitable for web construction. The softening behaviour typically seen in such silks, combined with a solid material structure rather than a discrete mesh, results in a greater spreading of damage that effectively enhances the system's fracture toughness. This is clearly an advantage for the protective role of cocoons, and is reminiscent of other biomaterials where mechanical robustness has been attributed to the formation of large plastic regions^{2,5}. The opposite is true for webs, where robustness arises from extreme localization of failure at sacrificial elements, with this behaviour enhanced by the stiffening of threads (Figs 1a and 2a).

The enhanced mechanical performance of the web relies on the integration of material and structure, which ultimately derives from the particular molecular structure of silk that features a composite of semi-amorphous protein and β -sheet nanocrystals. We suggest that web design principles might be considered in engineering, where current practice uses sacrificial elements solely to dissipate energy (for example, impact loading, seismic response). In spider webs, discrete sacrificial elements are instead a means to avoid potentially dangerous system-level loading and mitigate structural damage so that despite the small decrease in spider-web load capacity (Fig. 2b), the robustness of the structure overall is greatly enhanced (Fig. 3). This allows a spider to repair rather than rebuild completely, should failure occur. Such an engineering design could ignore the requirements for the magnitude of a potential load and allow local failure to occur, a design stipulation that requires the consideration of both material behaviour and structural architecture.

METHODS SUMMARY

The web consists of an arithmetic spiral⁷ supported by radial threads at regular intervals, constructed from two primary elements, radial threads and spiral threads⁷ (Fig. 1b), and is modelled using molecular dynamics procedures. We implement five material behaviours: (1) atomistically derived dragline silk behaviour (parameterized from molecular simulations of dragline silk^{16,17}) (Fig. 1a); (2) empirically parameterized dragline silk (from experimental data¹); (3) empirically parameterized viscid silk (from experimental data¹); (4) ideal linear elastic behaviour; (5) ideal elastic–perfectly plastic behaviour, incorporated in three arrangements in models A, B and C. We consider two types of application of loading, targeted (local) and global (wind) loading. To characterize the mechanical response of the web under targeted loading, a spring-load is imposed to a small section of the web and increased until failure is incurred (defined by the failure of loaded threads). Wind loading is applied via a constant drag force applied to all web threads. *In situ* experiments through simple mechanical assays are applied to

an orb web of the common European garden spider. We identify a web in its natural environment and deform radial and spiral threads using a mechanical applicator (a metal wire to load threads). During deformation we control the displacement and monitor images using a digital camera. For theoretical analysis we use quantized fracture mechanics³⁰, a theory that describes the failure of discrete structures such as a spider web and adapted here to incorporate the nonlinear stress–strain behaviour of silk. For a detailed description of the models see Methods and Supplementary Information.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 November 2010; accepted 25 November 2011.

- Gosline, J. M., Guerette, P. A., Ortlepp, C. S. & Savage, K. N. The mechanical design of spider silks: from fibroin sequence to mechanical function. *J. Exp. Biol.* **202**, 3295–3303 (1999).
- Gao, H., Ji, B., Jäger, I. L., Arzt, E. & Fratzl, P. Materials become insensitive to flaws at nanoscale: lessons from nature. *Proc. Natl Acad. Sci. USA* **100**, 5597–5600 (2003).
- Aizenberg, J. et al. Skeleton of *Euplectella* sp.: structural hierarchy from the nanoscale to the macroscale. *Science* **309**, 275–278 (2005).
- Vollrath, F. Spider webs and silks. *Sci. Am.* **266**, 70–76 (1992).
- Kamat, S., Su, X., Ballarín, R. & Heuer, A. H. Structural basis for the fracture toughness of the shell of the conch *Strombus gigas*. *Nature* **405**, 1036–1040 (2000).
- Lefèvre, T., Rousseau, M. E. & Pézolet, M. Protein secondary structure and orientation in silk as revealed by Raman spectromicroscopy. *Biophys. J.* **92**, 2885–2895 (2007).
- Vollrath, F. & Mohren, W. Spiral geometry in the garden spider's orb web. *Naturwissenschaften* **72**, 666–667 (1985).
- Vollrath, F. Silk evolution untangled. *Nature* **466**, 319 (2010).
- Agnarsson, I., Kuntner, M. & Blackledge, T. A. Bioprospecting finds the toughest biological material: extraordinary silk from a giant riverine orb spider. *PLoS ONE* **5**, e11234 (2010).
- Du, N. et al. Design of superior spider silk: from nanostructure to mechanical properties. *Biophys. J.* **91**, 4528–4535 (2006).
- Shao, Z. Z. & Vollrath, F. Materials: surprising strength of silkworm silk. *Nature* **418**, 741 (2002).
- Omenetto, F. G. & Kaplan, D. L. New opportunities for an ancient material. *Science* **329**, 528–531 (2010).
- Ko, K. K. et al. Engineering properties of spider silk. *Adv. Fibers Plastics Laminates Composites* **702**, 17–23 (2002).
- Rammensee, S., Slotta, U., Scheibel, T. & Bausch, A. R. Assembly mechanism of recombinant spider silk proteins. *Proc. Natl Acad. Sci. USA* **105**, 6590–6595 (2008).
- Vollrath, F., Holtet, T., Thøgersen, H. C. & Frische, S. Structural organization of spider silk. *Proc. R. Soc. Lond. B* **263**, 147–151 (1996).
- Keten, S., Xu, Z. P., Ihle, B. & Buehler, M. J. Nanoconfinement controls stiffness, strength and mechanical toughness of beta-sheet crystals in silk. *Nature Mater.* **9**, 359–367 (2010).
- Keten, S. & Buehler, M. J. Nanostructure and molecular mechanics of spider dragline silk protein assemblies. *J. R. Soc. Interf.* **7**, 1709–1721 (2010).
- Aoyanagi, Y. & Okumura, K. Simple model for the mechanics of spider webs. *Phys. Rev. Lett.* **104**, 038102 (2010).
- Ko, F. K. & Jovicic, J. Modeling of mechanical properties and structural design of spider web. *Biomacromolecules* **5**, 780–785 (2004).

20. Alam, M. S., Wahab, M. A. & Jenkins, C. H. Mechanics in naturally compliant structures. *Mech. Mater.* **39**, 145–160 (2007).
21. Alam, M. S. & Jenkins, C. H. Damage tolerance in naturally compliant structures. *Int. J. Damage Mech.* **14**, 365–384 (2005).
22. Foelix, R. F. *Biology of Spiders* 2nd edn (Oxford University Press/Georg Thieme, 1996).
23. Vollrath, F. Biology of spider silk. *Int. J. Biol. Macromol.* **24**, 81–88 (1999).
24. Termonia, Y. Molecular modeling of spider silk elasticity. *Macromolecules* **27**, 7378–7381 (1994).
25. Vepari, C. & Kaplan, D. L. Silk as a biomaterial. *Prog. Polym. Sci.* **32**, 991–1007 (2007).
26. Swanson, B. O., Blackledge, T. A. & Hayash, C. Y. Spider capture silk: performance implications of variation in an exceptional biomaterial. *J. Exp. Zool. A* **307A**, 654–666 (2007).
27. Swanson, B. O., Anderson, S. P., DiGiovine, C., Ross, R. N. & Dorsey, J. P. The evolution of complex biomaterial performance: the case of spider silk. *Integr. Comp. Biol.* **49**, 21–31 (2009).
28. Vollrath, F. & Selden, P. The role of behavior in the evolution of spiders, silks, and webs. *Annu. Rev. Ecol. Evol. Syst.* **38**, 819–846 (2007).
29. Keten, S. & Buehler, M. J. Atomistic model of the spider silk nanostructure. *Appl. Phys. Lett.* **96**, 153701 (2010).
30. Pugno, N. M. & Ruoff, R. S. Quantized fracture mechanics. *Phil. Mag.* **84**, 2829–2845 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported primarily by the Office of Naval Research (N000141010562) with additional support from the National Science Foundation (MRSEC DMR-0819762, the NSF-REU programme, as well as CMMI-0642545) and the Army Research Office (W911NF-09-1-0541 and W911NF-10-1-0127). Support from the MIT-Italy programme (MITOR) and a Robert A. Brown Presidential Fellowship is gratefully acknowledged. N.M.P. is supported by the METREGEN grant (2009–2012) “Metrology on a cellular and macromolecular scale for regenerative medicine”. An Ideas Starting Grant 2011 BIHSNAM on “Bio-inspired hierarchical super nanomaterials” was awarded to N.M.P. from the European Research Council, under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC grant (agreement number 279985). All simulations have been carried out at MIT’s Laboratory for Atomistic and Molecular Mechanics (LAMM). We acknowledge assistance from S. and E. Buehler in taking photographs of the spider web.

Author Contributions S.W.C. and M.J.B. designed the research and analysed the results. S.W.C. and A.T. developed the material models, performed the simulations, and conducted the simulation data analysis. M.J.B. performed the *in situ* experiments and analysed the results. N.M.P. contributed the theoretical analysis and predictions and analysed the results. S.W.C., M.J.B., A.T. and N.M.P. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.J.B. (mbuehler@mit.edu).

METHODS

Web geometry. Previous web models have implemented simplified versions of web geometry, most commonly in a concentric circle arrangement^{18–20}. Here we modelled a realistic orb web and approximated it by an arithmetic spiral⁷ defined by the polar equation $R(\theta) = \alpha \times \theta$, where spiral spacing is defined by $dR = 2\pi\alpha$, supported by radial threads at regular angular intervals ($d\theta = 45^\circ$). The basic web structure was constructed from two primary silk elements—radial threads and spiral threads⁷—combined with glue-like connections (Fig. 1b). The web was formed by particle-spring elements, with an equilibrium spacing of $r_0 = 0.01$ m. In nature, the construction of a stereotypical orb web includes the placement of framing threads that act as mooring and a structural foundation for the web⁷. The arrangement of such threads varies according to the anchoring points available to the spider and clearly has the strength necessary to bear the interior web, so we neglect it here. For a detailed description of the model, see Supplementary Information section 2.

Web models. We used web models based on combinations of five material behaviours: (1) atomistically derived dragline silk (parameterized from molecular simulations of dragline spider silk^{16,17}) (Fig. 1a); (2) empirically parameterized dragline silk (from experimental data¹); (3) empirically parameterized viscid silk (from experimental data¹); (4) ideal linear elastic behaviour; and (5) ideal elastic–perfectly plastic behaviour. To explore the differences between the theoretically derived silk models with experimentally measured silks, the materials were implemented in the following three web models.

Model A. We used atomistically derived dragline silk behaviour for both the radial and spiral threads (details in Supplementary Information section 3.1), to maintain independence from empirical data. Even though such a simple model formulation did not allow us to draw conclusions about phenomena pertaining to specific types of silk, it did enable us to understand universal, generic relationships between underlying molecular mechanisms, the resulting nonlinear properties of the material, and the failure behaviour of webs. The dragline radial behaviour was used for fitting corresponding linear elastic (model A', see Supplementary Information section 3.3) and elastic–perfectly plastic (model A'', see Supplementary Information section 3.4) models, as indicated in Fig. 2a.

Model B. We combined atomistically derived dragline silk for radial threads (see Supplementary Information section 3.1) with empirically parameterized viscid silk behaviour for spiral threads (see Supplementary Information section 3.2), to examine the effect of deviations in the stiffness of viscid silk (naturally more compliant than dragline silk). Because idealized behaviours (linear elastic or elastic–perfectly plastic) are parameterized on the basis of the radial response, there are no idealized iterations of model B.

Model C. This was a completely empirically parameterized web model, with empirically fitted dragline silk for the radial threads (described in Supplementary Information section 3.2) and empirically fitted viscid silk for spiral threads, for a realistic web representation tuned by experimental data¹. The empirical dragline behaviour is for fitting corresponding linear elastic (model C', see Supplementary Information section 3.3) and elastic–perfectly plastic behaviour (model C'', see Supplementary Information section 3.4).

The results of the empirically parameterized model are discussed in Supplementary Information section 5. The model framework used here can easily be adapted for other species of spiders, associated silk properties, and web geometries. Using a particle dynamics formulation (motivated by molecular dynamics), the total energy of the web system was defined as:

$$U_{\text{web}} = \sum_{\text{threads}} \phi_{\text{material}} \quad (1)$$

for the summation of the elastic potentials of all the silk threads, where ϕ_{material} refers to the constitutive energy expression of the specific material.

Atomistically derived dragline silk. To parameterize silk deformation behaviour, we use data from previous full atomistic simulations of major ampullate dragline spider silk^{16,17,32,33}, unaccounted for in previous web studies^{18–20}. The constitutive behaviour of dragline silk was formulated as:

$$\phi_{\text{dragline}}(r) = A_0 \times \begin{cases} \frac{1}{2} \frac{E_1}{r_0} (r - r_0)^2, & r \leq r_y \\ \frac{r_0}{\alpha} \exp\left[\frac{\alpha(r - r_y)}{r_0}\right] + \frac{1}{2} \frac{\beta}{r_0} (r - r_y)^2 + C_1(r - r_y) + C_3, & r_y \leq r < r_s \\ \frac{1}{2} \frac{E_2}{r_0} (r - r_s)^2 + C_2(r - r_s) + C_4, & r_s \leq r < r_b \\ 0, & r \geq r_b \end{cases} \quad (2)$$

See Supplementary Table 1 for all parameters.

Empirically parameterized silk. To assess the generality of the results obtained with our atomistically derived behaviour, we implemented empirically fitted material behaviours for models B and C'. The functional form of the empirically parameterized dragline (radial) silk was identical to that of the atomistically derived dragline silk (described by equation (2)). To represent the J-shaped viscid silk response measured in experimental studies, we used a combined linear and exponential function:

$$\phi_{\text{viscid}}(r) = A_0 \times \left(ar_0 \exp\left(\frac{r - r_0}{r}\right) + \frac{1}{2} br \frac{(r - 2r_0)}{r} + cr \right) \text{ for } r < r_b \quad (3)$$

We fitted the parameters in equations (2) and (3) to experiments on *Araneus diadematus*¹ for both dragline and viscid silk. See Supplementary Table 2 for all parameters.

Idealized material behaviours. For comparison, motivated by earlier studies³¹, we implemented a model that allowed us systematically to vary the nature of nonlinear behaviour, permitting cases of ideal linear elastic and ideal elastic–perfectly plastic (softening) behaviour, to develop general insight. The linear elastic behaviour was governed by:

$$\phi_{\text{linear}}(r) = \frac{1}{2} \left(\frac{E_{\text{linear}} A_0}{r_0} \right) (r - r_0)^2 \text{ for } r < r_b \quad (4)$$

while the elastic–perfectly plastic behaviour was governed by:

$$\phi_{\text{plastic}}(r) = \begin{cases} \frac{E_{\text{plastic}} A_0}{2r_0} (r - r_0)^2, & r_0 \leq r < r_y \\ \frac{E_{\text{plastic}} A_0}{2r_0} (r_y - r_0)^2 + \frac{E_{\text{plastic}} A_0}{r_0} (r_y - r_0)(r - r_y), & r_y \leq r < r_b \end{cases} \quad (5)$$

Both behaviours were parameterized to reflect either the ultimate stress and strain of atomistically derived dragline silk (in model A) or the ultimate stress and strain of empirically parameterized dragline silk (in model C) to provide a comparison between material laws and web performance. See Supplementary Tables 3 and 4 for all parameters.

Loading conditions. We considered two types of loading, targeted (local) and global (wind) loading. To characterize the mechanical response and robustness of the web under local load, a load was imposed on a small section of the web (see Supplementary Information section 4), representing, for example, a small piece of debris. The spring-load is increased until failure occurred (defined by the failure of all loaded threads). Load was imposed on a small section of the web in the out-of-plane direction, offset from the centre of the web. Proximity to the web centre maximizes the structural resistance of the entire web (as compared to loading the web periphery, for example), while the offset is used to apply the load to a known (chosen) radial thread to ease analysis. We determined the deflection of the web (out-of-plane) and applied force. We calculated the work needed to break individual threads by numerically integrating the force–displacement curves (see Supplementary Information section 6). To characterize the mechanical response under wind load (global), we applied a constant force to the entire web structure, derived from the equivalent drag force on a cylindrical wire (see Supplementary Information section 8). Loads for equivalent wind speeds of 0.5 to 70 m s^{−1} are applied (all models fail at 70 m s^{−1} winds).

In situ experimental studies. We carried out experiments on a physical web on the basis of mechanical assays applied to an orb web of the common European garden spider discovered in southern Germany. We identified a large spider web in its natural environment and ensured that the spider web was in use by a living spider. We deformed radial and spiral threads using a mechanical applicator, a small piece of wire that can effectively be used to pull on small structural features. During mechanical deformation of the web we controlled the displacement and monitored visual images of the web using a digital camera (results shown in Fig. 1e, f). A black plastic plate was placed behind the web to ensure that the web was clearly visible during the experiment for image acquisition.

Stress distribution. Normalized strain energy distributions were considered for radial threads just before and immediately after web fracture to calculate the average stress according to equations (2) to (5) (normalized with respect to maximum strain energy at ultimate failure). Spiral threads were not considered because most of the load (and thus elastic resistance) is carried by the radials in this load interval (see Supplementary Information section 9).

Theoretical analysis. We used quantized fracture mechanics³⁰, adapted here to incorporate the nonlinear material behaviour of silk using a generalized stress–strain (σ – ϵ) behaviour of $\sigma \propto \epsilon^\alpha$. The relative size of the damage zone after failure was given by:

$$\phi(x) = S^{2\alpha} \quad (6)$$

where S is a system-dependent constant reflective of specific material properties (such as fracture toughness), system geometry, and applied loading conditions.

The constant S describes the damage associated with the linear elastic behaviour when 2α tends to 1 and therefore $\varphi = S$. The fraction of surviving material after failure defines structural robustness:

$$\Omega(\alpha) = 1 - S^{2\alpha} \quad (7)$$

The parameter S was determined from the linear elastic response as the reference case, and constant for all variations in the stress–strain behaviour. The three material behaviours studied here (Fig. 2a), characteristic of silk, linear elastic and elastic–perfectly plastic behaviours, were reduced to general nonlinear stress–strain power

laws fitted by a single nonlinearity parameter α in the quantized fracture mechanics theory (Fig. 3). For details see Supplementary Information section 10.

31. Buehler, M. J. & Gao, H. Dynamical fracture instabilities due to local hyperelasticity at crack tips. *Nature* **439**, 307–310 (2006).
32. Nova, A., Keten, S., Pugno, N. M., Redaelli, A. & Buehler, M. J. Molecular and nanostructural mechanisms of deformation, strength and toughness of spider silk fibrils. *Nano Lett.* **10**, 2626–2634 (2010).
33. van Beek, J. D., Hess, S., Vollrath, F. & Meier, B. H. The molecular structure of spider dragline silk: folding and orientation of the protein backbone. *Proc. Natl Acad. Sci. USA* **99**, 10266–10271 (2002).

Decadal to monthly timescales of magma transfer and reservoir growth at a caldera volcano

T. H. Druitt^{1,2,3}, F. Costa⁴, E. Deloule⁵, M. Dungan⁶ & B. Scaillet⁷

Caldera-forming volcanic eruptions are low-frequency, high-impact events capable of discharging tens to thousands of cubic kilometres of magma explosively on timescales of hours to days, with devastating effects on local and global scales¹. Because no such eruption has been monitored during its long build-up phase, the precursor phenomena are not well understood. Geophysical signals obtained during recent episodes of unrest at calderas such as Yellowstone, USA, and Campi Flegrei, Italy, are difficult to interpret, and the conditions necessary for large eruptions are poorly constrained^{2,3}. Here we present a study of pre-eruptive magmatic processes and their timescales using chemically zoned crystals from the 'Minoan' caldera-forming eruption of Santorini volcano, Greece⁴, which occurred in the late 1600s BC. The results provide insights into how rapidly large silicic systems may pass from a quiescent state to one on the edge of eruption^{5,6}. Despite the large volume of erupted magma⁴ (40–60 cubic kilometres), and the 18,000-year gestation period between the Minoan eruption and the previous major eruption, most crystals in the Minoan magma record processes that occurred less than about 100 years before the eruption. Recharge of the magma reservoir by large volumes of silicic magma (and some mafic magma) occurred during the century before eruption, and mixing between different silicic magma batches was still taking place during the final months. Final assembly of large silicic magma reservoirs may occur on timescales that are geologically very short by comparison with the preceding repose period, with major growth phases immediately before eruption. These observations have implications for the monitoring of long-dormant, but potentially active, caldera systems.

Crystals in volcanic rocks provide records of magma reservoir processes and timescales before eruptions⁷. A crystal growing from a magmatic melt incorporates trace elements in quantities governed by thermodynamic and kinetic laws^{8,9}. If the crystal is subsequently mixed into another melt, trace elements that diffuse sufficiently fast will begin to re-equilibrate with the new host melt, generating intracrystalline diffusion gradients that can be used to obtain time information^{10–12} (Supplementary Fig. 1).

A novel technique for extracting multiple timescales from single crystals has been applied to plagioclase crystals in rhyodacitic (71–72 wt% SiO₂) pumices from the Minoan eruption. The pumices from all four phases of the eruption (phase 1, Plinian; phases 2 and 3, phreatomagmatic; phase 4, ignimbrite¹³) contain ~10 vol% crystals set in rhyolitic¹³ (73–74 wt% SiO₂) glass. Plagioclase is the dominant crystalline phase, the others being orthopyroxene, clinopyroxene, magnetite, ilmenite and apatite. It occurs as zoned crystals, whose compositional variations are dominated by charge-coupled solid solution between the endmembers, anorthite (An, CaAl₂Si₂O₈) and albite (NaAlSi₃O₈).

Two types of plagioclase have been recognized from examination of about 300 crystals (Fig. 1 and Supplementary Figs 2 and 3). A dominant

type (type 1) has euhedral to partly resorbed cores of An_{58–43} (58–43 mol% An) overgrown by variably developed, euhedral rims of An_{43–36} containing multiple dissolution surfaces (Fig. 1a–d). The cores and rims form compositionally distinct populations (Fig. 1g). In some crystals, the cores also contain euhedral to anhedral inner cores of An_{88–58} ('calcic inner cores'; Fig. 1a, c). A spectrum of type-1 crystals is observed, from those with thick rims (up to 100 μm) to those with thin rims to those in which the rim is absent (Fig. 1f). In rimmed type-1 crystals, the plagioclase in contact with the host glass is An_{~40}, whereas in rimless ones it is An_{~50} (that is, the core composition). Rimmed type-1 crystals occur in pumices from all four eruptive phases, but rimless ones have only been observed in pumices from phase 4. Type-2 crystals are very rare; they are reversely zoned, with cores of

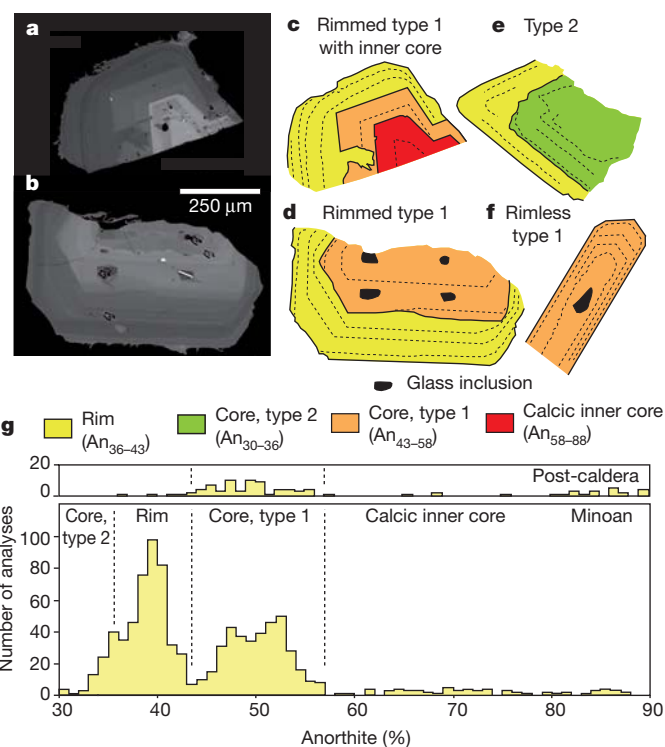


Figure 1 | Images and compositions of plagioclase crystals in Minoan pumice. a, b, Back-scattered electron images of two rimmed type-1 crystals (S82-30A 12 and S82-34B A2). c–f, Sketches of representative plagioclase crystals of the different types: S82-30A 12 (c); S82-34B A2 (d); S82-20D 76 (e); S82-20D C1 (f). More crystals are shown in Supplementary Fig. 2. g, Histogram of Minoan plagioclase compositions (bottom), based on 922 analyses, showing the different compositional groups. A histogram of plagioclase compositions from the post-caldera lavas^{16,17} (top) is shown for comparison.

¹Clermont Université, Université Blaise Pascal, Laboratoire Magmas et Volcans, BP 10448, F-63000 Clermont-Ferrand, France. ²CNRS, UMR 6524, LMV, F-63038 Clermont-Ferrand, France. ³IRD, R 163, LMV, F-63038 Clermont-Ferrand, France. ⁴Earth Observatory of Singapore, Nanyang Technological University, Singapore 639798. ⁵CRPG-CNRS, BP20, 54501 Vandœuvre les Nancy, France.

⁶Département de Minéralogie, Université de Genève, Rue des Maraichers, 13 CH-1205 Genève, Switzerland. ⁷ISTO-UMR 6113 CNRS/Université d'Orléans/Université François Rabelais de Tours, 1A rue de la Ferrollerie, 45071 Orleans, France.

An_{36–30} mantled by rims identical to those of type 1 (Fig. 1e). The broad range of plagioclase compositions in Minoan pumices shows that the rhyodacite was the product of open-system magmatic processes involving multiple, compositionally diverse magmas.

We analysed trace elements in four rimmed type-1 crystals, one rimless type-1 crystal, some interstitial glasses and some inclusions of glass contained within the crystals (Fig. 2 and Supplementary Tables 1 and 2). Magnesium (Mg), strontium (Sr) and titanium (Ti) are particularly useful elements for characterizing coexisting melt compositions and mixing timescales because they partition differently between melt and plagioclase⁸ and diffuse at different rates. Mg diffuses faster than Sr (ref. 7); Ti probably diffuses slowly owing to its high charge. We used published An-dependent partition coefficients⁸ to invert melt trace-element contents to those of equivalent plagioclase (Fig. 2), and vice versa (Fig. 3). Rimmed type-1 crystals have core-to-rim gradients in all three elements (Fig. 2). Rim compositions record equilibrium with interstitial glass, but concentrations of Mg, Sr and Ti in the cores and calcic inner cores are significantly higher than those calculated to be in equilibrium with the glass (Fig. 2, green lines). The rimless type-1 crystal has concentrations of Mg and Sr throughout that are too high to have been in equilibrium with the glass. None of these crystals resided in the host melt long enough for any of these elements to reach total equilibrium with the host melt.

This observation is reinforced by a comparison of calculated Mg, Ti and Sr concentrations of melts in equilibrium with the different plagioclase zones (referred to as Mg_{melt}, Ti_{melt} and Sr_{melt}, respectively), with those of Santorini lavas, pumices and glasses. The latter represent an approximate liquid line of descent of the magmatic system (Fig. 3a, b). The calculated concentrations of Sr_{melt} and Ti_{melt} span the entire range of erupted compositions, whereas those of Mg_{melt} fall below the liquid line of descent, especially in the calcic inner cores. These crystals have retained near-primary values of Sr and Ti, but Mg seems to have been partly re-equilibrated by diffusion.

We used the Sr contents of the crystals to calculate the compositions of melts from which the different zones crystallized (Fig. 3). The rims of the type-1 and type-2 crystals formed from melts with 60–90 p.p.m. Sr (71–74 wt% SiO₂), which encompasses the range of whole-rock to glass compositions of the host rhyodacite. The cores of the type-1 crystals formed from melts with 110–170 p.p.m. Sr (dacite with 63–69 wt% SiO₂), and the cores of the type-2 crystals formed from melts with 40–60 p.p.m. Sr (rhyolite with 74–76 wt% SiO₂). We interpret

these observations as recording the mixing of a dacite (carrying plagioclase of the type-1 core composition) with a rhyolite (carrying plagioclase of the type-2 core composition) to generate a hybrid rhyodacite (followed by growth of the rim) that was erupted as the Minoan pumice. The rimless type-1 crystals were derived directly from the dacite, and had not resided in the hybrid rhyodacite long enough to grow rims. The calcic inner cores formed from basaltic andesite melts, with >200 p.p.m. Sr and <57 wt% SiO₂, that were mixed into the dacite at an earlier stage. The possibility that the crystal cores and inner cores are derived from older plutonic mush is not favoured on the balance of evidence (Supplementary Discussion).

The times elapsed between the different mixing events and eruption have been calculated by modelling the partial relaxation of Mg concentrations in different parts of the crystals¹² (Fig. 2 and Table 1). We first modelled the entire crystal ('one-step model'), which assumes that diffusion started only after the entire crystal had grown. Model results for the elapsed time ranged from 1 to 60 yr, but good fits could not be obtained. We therefore approximated the more realistic situation of simultaneous crystal growth and diffusion by modelling the crystals incrementally zone by zone ('two-step model'). In this manner, we obtained better overall fits to the data (Fig. 2), from which we infer that (1) a first magma mixing episode recorded by incorporation of the calcic inner cores of the type-1 crystals into the dacite occurred no more than ~100 yr before eruption; (2) a second mixing episode, or series of episodes, involving the dacite and the rhyolite occurred less than ~10 yr before eruption; and (3) the rimless type-1 crystals cannot have resided in the Minoan magma more than a few months without some diffusive loss of Mg from the rim. We have no constraints on the ages of the cores of the rare type-2 crystals. These estimates are accurate to better than one order of magnitude (Methods). The Mg-derived times were tested against the diffusive equilibration of Sr using the comprehensive diffusion data for this element (Fig. 2). Taking the measured profiles of Sr as being close to initial values, diffusion times an order of magnitude longer than those obtained from Mg would significantly blur the observed tight correlation between Sr and An contents (Fig. 3c), and are ruled out.

Our reconstruction of events, which is justified more fully in Supplementary Discussion, is as follows. During the century before eruption, dacitic magma containing crystals of calcic plagioclase began to ascend from a deep storage zone beneath Santorini, possibly pushed by injections of mafic magma (Supplementary Fig. 4). Once at shallow

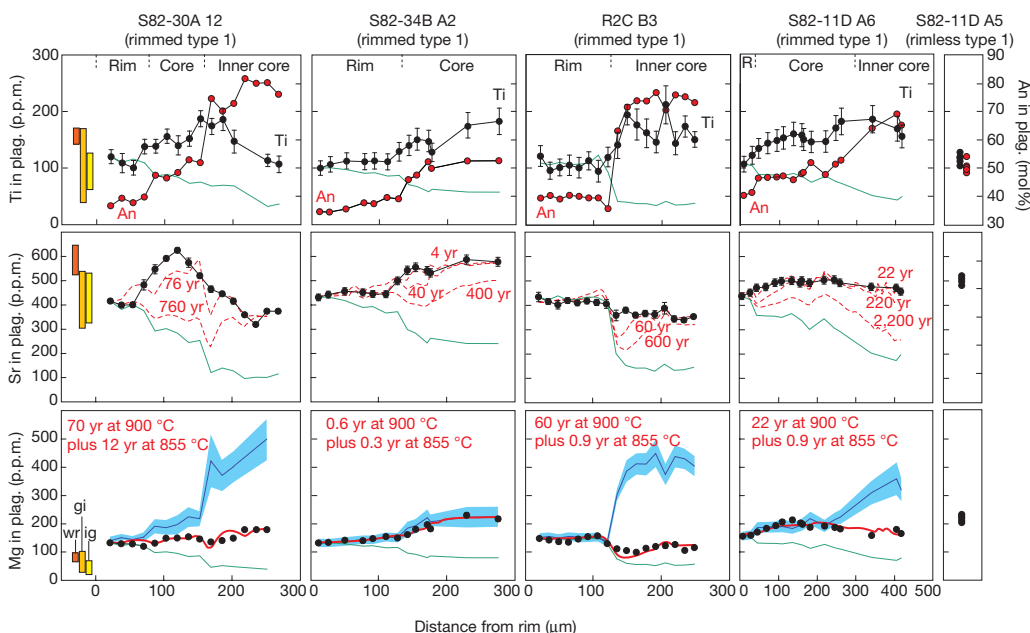


Figure 2 | Concentration-distance profiles of An (red dots) and of Sr, Ti and Mg (black dots) in Minoan plagioclase crystals. Equilibrium profiles for Sr, Ti and Mg (green lines) were calculated using published partition functions⁸ at 855 °C. Calculated plagioclase compositions in equilibrium with interstitial glasses (ig), glass inclusions (gi) and whole-rock (wr) pumice compositions¹³ are also shown. The blue lines in the Mg plots are the initial Mg profiles (with 1σ range). The red solid lines in the Mg plots are the best-fit two-step Mg diffusion models (time shown in years). The red dotted lines in the Sr plots are diffusion models (one-step models; 900 °C; time shown in years) in which we assume that the measured profile is close to the initial condition, as suggested by Fig. 3a. Error bars, 1σ.

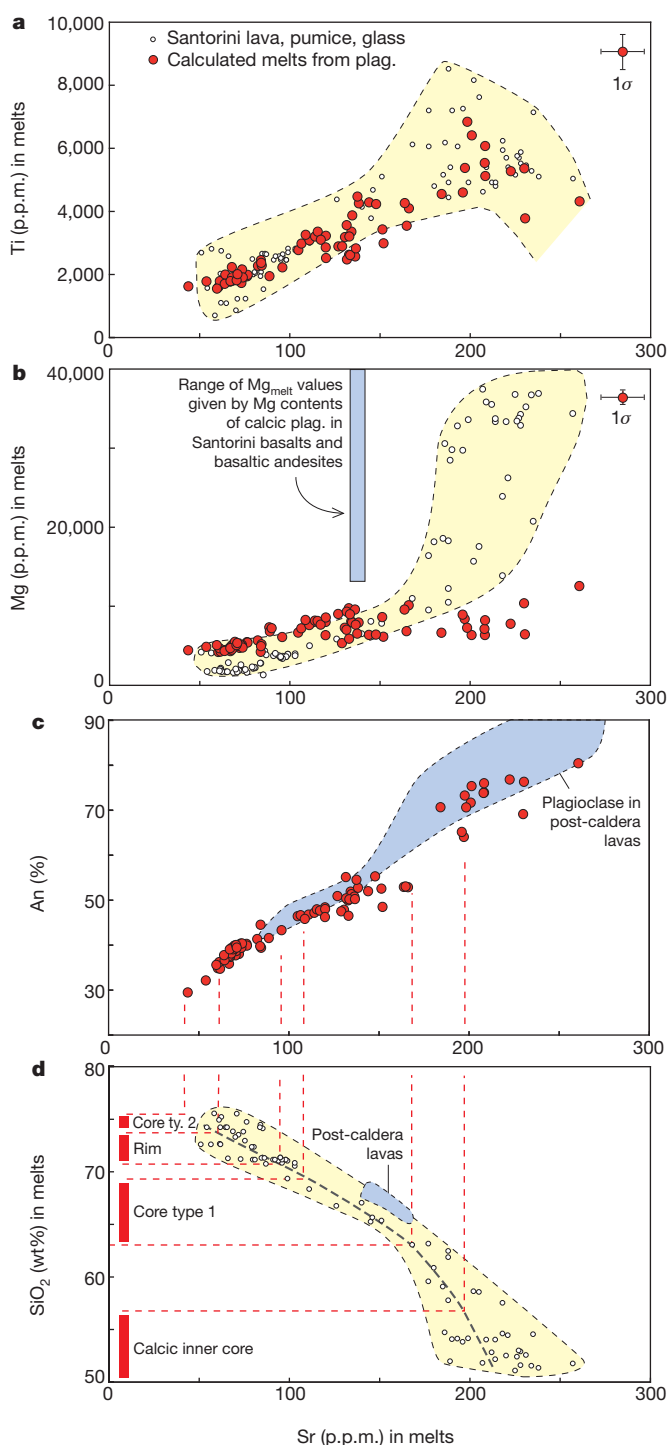


Figure 3 | Melt compositions calculated by inversion of plagioclase trace-element compositions. **a**, **b**, Ti, Mg and Sr melt contents calculated using the plagioclase–liquid partition functions of ref. 8. Temperature was obtained from $T = 855\text{ }^{\circ}\text{C} + 200\text{ }^{\circ}\text{C} \times (\% \text{An} - \% \text{An}_{40}) / (\% \text{An}_{80} - \% \text{An}_{40})$, which assumes that An_{40} plagioclase crystallized at $855\text{ }^{\circ}\text{C}$ and that An_{80} crystallized at $1,055\text{ }^{\circ}\text{C}$, the temperature of Santorini basaltic andesite. The values are compared with the compositional fields of Santorini cycle-2 lavas¹⁷, Minoan pumices and Minoan glasses. The range of Mg_{melt} values calculated in the same way from the Mg contents of calcic ($\text{An}_{>80}$) plagioclase in Santorini basalts and basaltic andesites are much higher (14,000–47,000 p.p.m.) than those in the calcic inner cores of Minoan crystals, showing that the latter have lost Mg by diffusion. **c**, **d**, Use of calculated melt Sr contents from plagioclase allows estimation of the SiO_2 contents of the melts from which each plagioclase compositional group crystallized. The compositional field of plagioclases in the post-caldera lavas is taken from ref. 32.

levels, the dacite intersected one or more extant bodies of rhyolite situated at a depth of a few kilometres¹⁴. The cores of the type-1 crystals grew during the ascent of the dacite and/or its ponding at the base of the reservoir. Then, starting about a decade before eruption, the dacite mixed with the rhyolite to form rim overgrowths on the type-1 crystals. Mixing caused resorption of most crystals initially present in the rhyolite, and formed the hybrid rhyodacite that was subsequently erupted. Successive waves of recharge and heating of the reservoir are recorded as multiple dissolution surfaces within crystal rims. Dacite was still being mixed into the shallow reservoir less than a few months before eruption, supplying the rimless type-1 crystals. Efficient mixing of the two silicic magmas on these timescales is compatible with fluid dynamics principles (Supplementary Discussion). Andesite and basaltic andesite were also injected at a late stage, accounting for small quantities of quenched enclaves discharged during the first eruptive phase¹³. The plagioclase crystals in the rhyodacite have lower Sr isotopic ratios than their host glass¹⁵, so the resident rhyolite may have been more radiogenic than the dacite owing to crustal contamination. The dacite may have been an advance batch of the magmas that have erupted over the last 2,000 yr to form the post-caldera Kameni Islands¹³. This interpretation is based on similarities in An and Sr contents between the cores of type-1 plagioclases and those of plagioclases in the post-caldera lavas^{16,17} (Figs 1g and 3c).

Calculations based on chemical mass balance and magma crystallinity show that the dacitic component accounted for at least 15% of the erupted magma (Supplementary Discussion). Assembly of the Minoan reservoir during the 18,000 yr since the previous major eruption of Santorini terminated in a growth spurt involving injection of at least several cubic kilometres of dacitic magma, and an unknown volume of more-mafic magma, in about a century. This implies a transient recharge flux ($>0.05 \text{ km}^3 \text{ yr}^{-1}$) much higher than the long-term magma production rate at the volcano¹³ ($\sim 0.001 \text{ km}^3 \text{ yr}^{-1}$). Rapid and sustained supply of compositionally diverse magmas eventually led to overturn and homogenization of the different magma batches, and eruption followed about a decade later.

Accommodation of such a large recharge volume purely by roof uplift could have caused great inflation of the volcano during the century preceding eruption. A maximum estimate of this is given by the elevation of a rigid, piston-like roof block with the same area as the caldera (several tens of metres of uplift at a mean rate of $\sim 1 \text{ m yr}^{-1}$; broader doming of the volcano¹⁸, or consideration of rock and magma compressibilities, would give lower values). Although this is possible, such high uplift values have not been observed, or inferred, before past caldera-forming eruptions such as that of Mt Pinatubo in 1991. However, exhumed granitoid plutons reveal that injection of silicic magma at depths of $>3 \text{ km}$ is often accommodated by lowering of the intrusion floor due to down-sagging of underlying crystal-rich magmatic mushes and crustal layers^{5,19–21}. Extraction of magma from deep storage regions can result in compaction of the crustal column beneath the volcano. Had foundering beneath the Minoan reservoir taken place rapidly enough, precursory uplift might have been greatly reduced. Subsidence rates of several decimetres per year beneath growing magma reservoirs are compatible with theoretical models^{5,20}.

Our study shows that the reactivation, growth and final assembly of large silicic magma bodies before eruption can take place on decadal to monthly timescales that are very short by comparison with the lengths of the preceding repose periods. It is consistent with evidence for voluminous, late-stage recharge on similar timescales before caldera eruptions elsewhere^{11,22–24}. High rates of magma supply sustained over many decades may be necessary to ‘prime’ such bodies for eruption, by generating overpressures high enough to propagate dykes of viscous magma to the surface and initiate discharge^{25,26}. Long-term monitoring of large, dormant caldera systems^{27–29}, even in remote areas of the world, is essential if late-stage growth spurts of shallow magma reservoirs are to be detected well in advance of caldera-forming eruptions.

Table 1 | Mixing-to-eruption times obtained by diffusion modelling of Mg in Minoan plagioclase

Pumice sample	Eruption phase	Crystal (type)	Time (yr), one-step model	Time (yr), two-step model				
				CIC-dm 900 °C	CIC-C-R-rm 855 °C	CIC-R-rm 855 °C	C-rm 855 °C	C-R-rm 855 °C
S82-30A	1	12 (rimmed 1)	47 ⁺³⁸ ₋₂₀	70 ⁺⁵⁶ ₋₃₀	12 ⁺¹⁰ ₋₅	—	—	—
S82-34B	2	A2 (rimmed 1)	0.8 ^{+0.6} _{-0.3}	—	—	—	0.6 ^{+0.5} _{-0.3}	0.3 ^{+0.3} _{-0.1}
R2C	3	B3 (rimmed 1)	60 ⁺⁴⁸ ₋₂₆	60 ⁺⁴⁸ ₋₂₆	—	<0.9 ^{+0.8} _{-0.4}	—	—
S82-11D	4	A6 (rimmed 1)	24 ⁺²⁰ ₋₁₀	22 ⁺¹⁷ ₋₉	0.9 ^{+0.8} _{-0.4}	—	—	—
S82-11D	4	A5 (rimless 1)	—	—	—	—	<0.3 ^{+0.3} _{-0.1}	—

The calculations were done at the temperatures shown, which correspond to those of dacitic and rhyodacitic melts. The quoted errors correspond to a temperature uncertainty of +25 °C in the calculated time. CIC, calcic inner core; C, core; R, rim; dm, dacitic melt; rm, rhyodacitic melt. The different crystals are shown in Supplementary Fig. 2.

METHODS SUMMARY

The zoning textures of about 300 crystals from eight pumices (from all four eruptive phases) were studied using a combination of back-scattered electron microscopy and Normarski differential interference contrast microscopy. Major-element compositions of minerals and glasses were analysed by electron microprobe, and trace elements by secondary ion mass spectrometry. The pre-eruptive temperature of the Minoan rhyodacite (853 ± 12 °C) was calculated using analyses of touching magnetite-ilmenite pairs in pumices from all eruptive phases. Trace-element diffusion modelling was carried out in one dimension using a forward modelling approach in which the diffusion equation¹² was solved numerically by a finite-difference scheme, and visual best fits were made between modelled and observed concentration profiles. The An-dependent diffusion coefficient for Mg was taken from ref. 12, which is based on the experimental data of ref. 30, and that for Sr was taken from ref. 31. Initial profiles for Mg were calculated using the Sr–An correlation of Fig. 3c and the Mg–Sr liquid line of descent (Fig. 3b). The boundary conditions were determined by the measured concentrations at the crystal rims or at other major boundary zones. The calculated times are probably overestimates because some of the cores have re-equilibrated significantly and, hence, the effect of diffusion in multiple dimensions could have been significant⁷. However, the main errors associated with the calculations arise from uncertainties in temperature and diffusion coefficients. The modelling was carried out at either 855 or 900 °C, depending on the melt involved (Table 1), but we considered a nominal uncertainty of ± 25 °C, which when propagated into the time estimates leads to relative errors of 45–75% (Table 1). Effects of uncertainties in the diffusion parameters for Mg are discussed in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 May; accepted 4 November 2011.

- Miller, C. F. & Wark, D. A. Supervolcanoes and their explosive supereruptions. *Elements* **4**, 11–15 (2008).
- Newhall, C. G. & Dzurizin, D. *Historical Unrest at Large Calderas of the World Vols 1 and 2* (Bull. US Geol. Surv. 1855, USGS, 1988).
- Gottsmann, J. & Marti, J. (eds) *Caldera Volcanism: Analysis, Modelling and Response* (Dev. Volcanol. 10, Elsevier, 2008).
- Sigurdsson, H. & Carey, S. and 12 others. Marine investigations of Greece's Santorini volcanic field. *Trans. Am. Geophys. Union* **87**, 337–342 (2006).
- Bachmann, O. & Bergantz, G. W. On the origin of crystal-poor rhyolites: extracted from batholithic crystal mushes. *J. Petrol.* **45**, 1565–1582 (2004).
- Burgisser, A. & Bergantz, G. W. A rapid mechanism to remobilize and homogenize highly crystalline magma bodies. *Nature* **471**, 212–215 (2011).
- Costa, F., Dohmen, R. & Chakraborty, S. Timescales of magmatic processes from modeling the zoning patterns of crystals. *Rev. Mineral. Geochem.* **69**, 545–594 (2008).
- Bindeman, I. N., Davis, A. M. & Drake, M. J. Ion microprobe study of plagioclase-basalt partition experiments at natural concentration levels of trace elements. *Geochim. Cosmochim. Acta* **62**, 1175–1193 (1998).
- Blundy, J. & Wood, B. Crystal-chemical control on the partitioning of Sr and Ba between plagioclase feldspar, silicate melts, and hydrothermal solutions. *Geochim. Cosmochim. Acta* **55**, 193–209 (1991).
- Zellmer, G. F., Blake, S., Vance, D., Hawkesworth, C. & Turner, S. Plagioclase residence times at two island arc volcanoes (Kameni Islands, Santorini, and Soufriere, St Vincent) determined by Sr diffusion systematics. *Contrib. Mineral. Petrol.* **136**, 345–357 (1999).
- Morgan, D. J. *et al.* Magma chamber recharge at Vesuvius in the century prior to the eruption of A.D. 79. *Geology* **34**, 845–848 (2006).
- Costa, F., Chakraborty, S. & Dohmen, R. Diffusion coupling between trace and major elements and a model for calculation of magma residence time using plagioclase. *Geochim. Cosmochim. Acta* **67**, 2189–2200 (2003).
- Druitt, T. H. *et al.* Santorini Volcano (J. Geol. Soc. Lond. Mem. 19, Geological Society, 1999).

- Cottrell, E., Gardner, J. E. & Rutherford, M. J. Petrologic and experimental evidence for the movement and heating of the pre-eruptive Minoan rhyodacite (Santorini, Greece). *Contrib. Mineral. Petrol.* **135**, 315–331 (1999).
- Martin, V., Davidson, J., Morgan, D. & Jerram, D. Using the Sr isotope compositions of feldspars and glass to distinguish magma system components and dynamics. *Geology* **38**, 539–542 (2010).
- Francalanci, L. *et al.* in *The European Laboratory Volcanoes 175–186* (Official Publ. Eur. Comm., 1998).
- Huijsmans, J. *Calc-Alkaline Lavas from the Volcanic Complex of Santorini, Aegean Sea, Greece* (Geologica Ultraiectina 41, Inst. Aardwetenschappen Rijksuniversiteit Utrecht, 1985).
- Aizawa, K., Acocella, V. & Yoshida, T. How the development of magma chambers affects collapse calderas: insights from an overview. *Spec. Publ. Geol. Soc. (Lond.)* **269**, 65–81 (2006).
- Wiebe, R. A. & Collins, W. J. Depositional features and stratigraphic sections in granitic plutons: implications for the emplacement and crystallization of granitic magma. *J. Struct. Geol.* **20**, 1273–1289 (1998).
- Cruden, A. R. On the emplacement of tabular granites. *J. Geol. Soc. Lond.* **155**, 853–862 (1998).
- Grocott, J., Arévalo, C., Welkner, D. & Cruden, A. Fault-assisted vertical pluton growth: Coastal Cordillera, north Chilean Andes. *J. Geol. Soc. Lond.* **166**, 295–301 (2009).
- Wark, D. A., Hildreth, W., Spear, F. S., Cherniak, D. J. & Watson, E. B. Pre-eruption recharge of the Bishop magma system. *Geology* **35**, 235–238 (2007).
- Saunders, K. E., Morgan, D. J., Baker, J. A. & Wysoczanski, R. J. The magmatic evolution of the Whakamaru supereruption, New Zealand, constrained by a microanalytical study of plagioclase and quartz. *J. Petrol.* **51**, 2465–2488 (2010).
- de Silva, S., Salas, G. & Schubring, S. Triggering explosive eruptions: the case for silicic magma recharge at Huaynaputina, southern Peru. *Geology* **36**, 387–390 (2008).
- McLeod, P. & Tait, S. R. The growth of dykes from magma chambers. *J. Volcanol. Geotherm. Res.* **92**, 231–245 (1999).
- Jellinek, A. M. & DePaulo, D. J. A model for the origin of large silicic magma chambers: precursors of caldera-forming eruptions. *Bull. Volcanol.* **65**, 363–381 (2003).
- Gottsmann, J. & Battaglia, M. Deciphering causes of unrest at explosive collapse calderas: Recent advances and future challenges of joint time-lapse gravimetric and ground deformation studies. *Dev. Volcanol.* **10**, 417–446 (2008).
- Hill, D. P. Unrest in Long Valley caldera, California, 1978–2004. *Spec. Publ. Geol. Soc. (Lond.)* **269**, 1–24 (2006).
- Dzurizin, D., Yamashita, K. M. & Kleinman, J. W. Mechanisms of crustal uplift and subsidence at the Yellowstone caldera, Wyoming. *Bull. Volcanol.* **56**, 261–270 (1994).
- LaTourrette, T. & Wasserbourg, G. J. Mg diffusion in anorthite: implications for the formation of early solar system planetesimals. *Earth Planet. Sci. Lett.* **158**, 91–108 (1998).
- Gilletti, B. J. & Casserly, J. E. D. Strontium diffusion kinetics in plagioclase feldspars. *Geochim. Cosmochim. Acta* **58**, 3785–3793 (1994).
- Santo, A. P. Magmatic evolution processes as recorded in plagioclase phenocrysts of Nea Kameni rocks (Santorini Volcano, Greece). *Dev. Volcanol.* **7**, 139–160 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was funded partly by the French Agence National de Recherche (ANR STOMIXAN, contract no. ANR-08CEAO080, to B.S.). We are grateful to R. Armstrong, P. Crançon and R. Girardin for their contributions during the early stages of this study, and to J. Blundy and M. Reid for reviews. This is Laboratory of Excellence ClerVolc contribution no. 1.

Author Contributions T.H.D. defined the project strategy, analysed the data and wrote the first draft of the manuscript, which was then revised by all the authors. E.D., M.D. and T.H.D. made the trace-element analyses, F.C. did the diffusion modelling and B.S. performed the fluid dynamic calculations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.H.D. (tdruitt@opgc.univ-bpclermont.fr).

METHODS

Sample preparation and analysis. Plagioclase crystals (250–400- μm and 400–700- μm size fractions) from eight pumices from all four eruptive phases¹³ were mounted in epoxy resin, mostly on (010), then polished. The zoning textures of about 300 crystals were studied using a combination of back-scattered electron microscopy and Normarski differential interference contrast microscopy. Sectioning through the centres of the crystals was ensured by visual observation using a binocular microscope and by rejection of zoning images that were evidently oblique cuts. Major-element compositions of plagioclase were analysed using a Cameca SX100 electron microprobe at an accelerating voltage of 15–20 kV and beam current of 8–15 nA. The pre-eruptive temperature of the Minoan rhyodacite ($853 \pm 12^\circ\text{C}$) was calculated using analyses of touching magnetite–ilmenite pairs in pumices from all eruptive phases, the recalculation procedure of ref. 33, and the thermometer of ref. 34.

Abundances of a number of isotopes as far up the periodic table as ^{140}Ce , including ^{24}Mg , ^{47}Ti and ^{88}Sr , were analysed using two ion probes: a Cameca IMS 4f instrument at the University of Edinburgh (4.5–5-nA primary O^- beam) and a Cameca IMS 1280 instrument at the CRPG, Nancy (14–20-nA primary O^- beam), using beam diameters of 10–15 μm and calibrated against glass standards NBS 6104 and ML3B-G. Secondary ions were energy-filtered using band passes of 100 and 35 V, offsets of 100 and 60 V, and mass resolutions of 300 and 5,000 on the IMS 4f and the IMS 1280, respectively, following a short pre-sputtering period. Only high-energy ions were measured, to reduce the matrix effect on relative ion yields³⁵. The calibration line for Sr, including glass standards StHs, KL2G, TIG BHVO and BCR2G, and CRPG reference albite and plagioclase, shows no matrix effect for a large range of glass and mineral compositions. Isotope concentrations were expressed as ratios relative to ^{30}Si , and were then corrected using electron microprobe analyses of Si contents adjacent to the ion probe pits. The relative ionization yields were determined on glass standards. Internal precisions (mean relative deviation) on the ion/ ^{30}Si ratios for each spot analysis were typically better than 5% for Mg, 5% for Sr and 10% for Ti. Values for all elements on the IMS 4f were systematically slightly lower than those on the IMS 1280, and were raised using a small multiplication factor.

Selected profiles were also analysed for Mg using a Cameca SX100 electron microprobe (accelerating voltage, 15 kV; beam current, 40 nA; total counting time per spot, 50 s), and using a LA-ICP-MS system consisting of a 193-nm excimer (ArF) laser-ablation system (M50, Resonetics) coupled to an Agilent 7500CS ICP-MS (10- μm beam size; 6-mJ beam energy; 1-Hz repetition rate). Results from the three techniques agreed to within 25%.

Crystal–melt trace-element partitioning. We used the An-dependent partition functions in ref. 8 for Mg, Sr and Ti. The Sr partition function in ref. 9 gives results within a few per cent of that in ref. 8. The partition functions for Mg, Sr and Ti in ref. 8 seem to work well for Santorini compositions. There is good agreement between the contents of Sr and Ti in the crystal rims and the plagioclase compositions calculated to be in equilibrium with the interstitial glasses (Fig. 2). Agreement for Mg is not as good, although the disparity is small in comparison with the overall range of this element in Santorini plagioclases and melts (Figs 2 and 3b). Strontium and Ti contents of the calcic inner cores of Minoan plagioclase crystals agree with those predicted in ref. 8 to be in equilibrium with Santorini basaltic and basaltic andesite melts (Fig. 3a). Magnesium contents of calcic plagioclase occurring in Santorini basalts and basaltic andesites^{36–38} agree with the Mg contents of those rocks (Fig. 3b) using the Mg partition coefficient in ref. 8.

Diffusion modelling. Diffusion modelling was carried out using the method of ref. 12 with distance increments of 3.5–5.3 μm . We used the Mg diffusion coefficient in plagioclase in ref. 12, which is based on the experimental data in ref. 30:

$$D_{\text{Mg}} = \left[2.92 \times 10^{-4.1X_{\text{An}} - 3.1} \exp\left(\frac{-266,000}{RT}\right) \right] \quad (1)$$

The diffusion coefficient for Sr was taken from ref. 31:

$$D_{\text{Sr}} = \left[2.92 \times 10^{-4.1X_{\text{An}} - 4.08} \exp\left(\frac{-276,000}{RT}\right) \right] \quad (2)$$

In these expressions, D is the element diffusion coefficient in $\text{m}^2 \text{s}^{-1}$, X_{An} is the mole fraction of anorthite, T is the temperature in kelvin and R is the ideal gas constant. According to equations (1) and (2), $D_{\text{Mg}} > D_{\text{Sr}}$ at all temperatures in plagioclase of a given An content. We used the equation for trace-element diffusion in plagioclase derived in ref. 12, which takes into account the effect of An content on equilibrium profile shapes and diffusion coefficients:

$$\frac{\partial C_i}{\partial t} = \left(D_i \frac{\partial^2 C_i}{\partial x^2} + \frac{\partial C_i}{\partial x} \frac{\partial D_i}{\partial x} \right) - \frac{A}{RT} \left(D_i \frac{\partial C_i}{\partial x} \frac{\partial X_{\text{An}}}{\partial x} + C_i \frac{\partial D_i}{\partial x} \frac{\partial X_{\text{An}}}{\partial x} + D_i C_i \frac{\partial^2 X_{\text{An}}}{\partial x^2} \right)$$

Here A is a thermodynamic factor that relates the dependence of the partition coefficient on the An content⁸, x is distance, C is concentration and t is time.

We used two temperatures in the modelling: 900°C for models in which the host melt was dacitic and 855°C for those in which it was rhyodacitic (Table 1). These temperatures are constrained by the temperature (855°C) of the magma on eruption (from touching pairs of Fe–Ti oxides), and are consistent with the formula used for plagioclase–melt element partitioning (Fig. 3).

The equilibrium distributions of Mg, Sr and Ti within the plagioclase crystals were obtained as follows. We first calculated the plagioclase–melt partition coefficients using the measured An contents and the equations in ref. 8. Using the partition coefficients, and the measured trace-element concentrations at the crystal rims, we then calculated a melt composition in equilibrium with the rim. Finally, the composition of the entire plagioclase crystal in equilibrium with the calculated melt (and crystal rim) was calculated using the partition coefficient for each portion of the crystal.

The initial Mg profile was estimated using the Sr–An correlation of Fig. 3c and the Mg–Sr liquid line of descent (Fig. 3b). The relationship between the initial Mg and An contents used in the modelling is $[\text{Mg}] = 48.2 \exp(2.9X_{\text{An}})$, where the concentration of Mg (left-hand side) is measured in p.p.m. This approach is justified because the times involved in the processes are short enough not to have affected Sr concentrations in any significant manner. The uncertainty in the initial Mg content shown in Fig. 2 was estimated by error propagation. For the boundary conditions, we used measured concentrations at the rims of the crystals or at the appropriate boundary between two crystal zones in the multiple-step models.

In the modelling, we explored two endmember possibilities. The first is to model the entire crystal, including the rim, core and calcic inner core, all at once (one-step model) at 900°C . This is equivalent to assuming that diffusion started only after the crystal was fully grown, and is the approach of all previously published models using plagioclase. This should not be a bad approximation if the diffusion time is relatively long and if growth rates are relatively high. The second possibility is to model the crystal piece by piece; for example, first model equilibration of the calcic inner core with dacitic melt at 900°C , then add the core, rim and rhyodacitic melt at 855°C , and continue diffusing (two-step model). This is a first-order approximation for modelling simultaneous crystal growth and diffusion. We tried both approaches, and in most cases there were no significant differences in total times, but the fits to the data were better when we used the two-step approach. This also gave us additional information about the timescales of different processes recorded in different parts of the crystals. To our knowledge this is the first time that plagioclase zoning has been used to provide chronologies of multiple events from single crystals. We note that to model some crystals successfully, we needed to use a condition where Mg in the crystal was already equilibrating with the host melt before any subsequent overgrowth by a new composition.

We illustrate the two approaches using crystal S82-30A 12 (Supplementary Fig. 5). We were not able to obtain a good fit to the measured Mg profiles using a one-step model, which gave an approximate time of 47 yr. We therefore modelled the diffusion using the two-step approach. We first did an equilibration of the calcic inner core with dacitic melt at 900°C , and obtained a time of 70 yr (step 1 of Supplementary Fig. 4b). We then added the core and rim over the calcic inner core and continued to diffuse in rhyodacitic melt at 855°C , and obtained a further time of 12 yr (step 2 of Supplementary Fig. 5b). The total time recorded by the crystal is therefore 82 yr, with a fit that is clearly better than in a one-step approach. We also tried more detailed models involving, for example, three steps (inner core/melt, inner core/core/melt and inner core/core/rim/melt), but the results were virtually identical to those of the two-step model.

In the case of the rimless type-1 crystal (S82-11D A5), the composition of which is out of equilibrium with the host rhyodacitic melt, we modelled the time necessary for Mg and Sr in the outer 25 μm of the crystal to be significantly modified by diffusive re-equilibration. The times at 855°C were 0.3 yr for Mg (Table 1) and 12 yr for Sr. The crystal therefore cannot have resided in the melt for longer than 0.3 yr.

The times obtained from Mg diffusion calculations are subject to various types of uncertainty⁷. First, the temperatures obtained from geothermometers have errors of about $\pm 25^\circ\text{C}$, which when propagated into the time estimates lead to the relative uncertainties given in Table 1. Second, the diffusion calculations were carried out in one dimension, which has been shown to overestimate actual (three-dimensional) diffusion time⁷ when the extent of diffusion is as significant as that shown by some of the inner cores of our crystals. Third, uncertainty arises from experimentally determined diffusion coefficients. Magnesium diffusion in

plagioclase was measured in An₉₅ and at 1,200–1,400 °C (ref. 30), and the measurements are precise to within a factor of two. Following ref. 12, equation (1) assumes that the activation energy and pre-exponential factor determined for D_{Mg} at higher temperatures applies at 900 °C. Down-temperature extrapolation of diffusion data is common practice and a necessity in almost all studies that use experimentally determined diffusion coefficients to model natural processes^{39,40}. This is because the experimental times required to produce significant profiles to determine diffusion coefficients are impracticably long (but see ref. 41). The extrapolation is not a significant problem if there is no change of the diffusion mechanism (which would imply a change in the activation energy) at lower temperatures. There have been many determinations of other cation diffusivities⁴² (for example those of Sr, Li and Ca) in plagioclase feldspars over wide ranges of temperature (including 900 °C), and these do not show kinks in activation energy.

There is also uncertainty in the An dependence in equation (1), which assumes that the An dependence of Mg diffusion is the same as that of Sr (ref. 12). This assumption was used in ref. 12 because only certain dependencies are possible if good fits to the natural data are to be obtained. In addition to this observation, we have here estimated an uncertainty due to the An dependence term in equation (1) by combining natural data and the well-calibrated Sr diffusivities. The data in Table 1 and Fig. 2, as well as those published elsewhere^{7,43,44}, show that the degree of equilibration of Mg is always higher than that of Sr, irrespective of An content. This means that $D_{Mg} > D_{Sr}$ not only for An₉₅ (ref. 30) but for all An contents. Thus, the dependence of D_{Mg} on An cannot be too low because otherwise $D_{Sr} > D_{Mg}$ at low An content. Using this constraint, and equation (2) for D_{Sr} , the minimum possible dependence on An (that is, the first exponent) in equation (1) has the form $-2.6X_{An} - 4.53$, which in a plot of D_{Mg} versus %An (Supplementary Fig. 6) gives a much lower slope than that given by equation (1). The effect that this has on the times ranges from an increase by a factor of two in the one-step fit for crystal R2C B3 (130 yr) to an increase by a factor of three for crystal 34BA2 (2.3 yr). We conclude that the uncertainty in this parameter does not increase the times in Table 1 by more than about a factor of three.

We also carried out diffusion modelling of Sr (for which the diffusion parameters are very well constrained³¹; Fig. 2), to gain first-order independent constraints on the timescales inferred from Mg. For this, we assumed that the observed profiles were near primary and carried out one-step diffusion models at 900 °C to see what times were required for significant perturbation of those profiles. The results (Fig. 2) show that crystal residence times longer than a couple of hundred years would

significantly blur the tight relationship between plagioclase Sr and An content (Fig. 3c), and are therefore excluded.

Finally, there is also an uncertainty in the determination of the best fits of the model and measured concentrations, which include the analytical errors. The fits were done visually because, given the complexity of the natural profiles, it does not make sense to implement least-squares routines. The uncertainty from this varies from crystal to crystal and ranges from less than 10% for crystals that have been largely equilibrated (for example by step one of most crystals) to up to 40% for step two, where concentrations vary by a small amount and may be close to the initial values (for example in crystal S82-34B A2).

33. Stormer, J. C. Jr. The effects of recalculation on estimates of temperature and oxygen fugacity from analyses of multicomponent iron-titanium oxides. *Am. Mineral.* **68**, 586–594 (1983).
34. Andersen, D. J. & Lindsley, D. H. New (and final!) models for the Ti-magnetite/ilmenite geothermometer and oxygen barometer. *Trans. Am. Geophys. Union* **66**, 416 (1985).
35. Hinton, R. W. Ion microprobe trace-element analysis of silicates: measurement of multi-element glasses. *Chem. Geol.* **83**, 11–25 (1990).
36. Vaggelli, G., Pellegrini, M., Vougioukalakis, G., Innocenti, S. & Francalanci, L. Highly Sr radiogenic tholeiitic magmas in the latest inter-Plinian activity of Santorini volcano, Greece. *J. Geophys. Res.* **114**, B06201 (2009).
37. Conticelli, S., Francalanci, L., Santo, A. P. & Petrone, C. in *The European Laboratory Volcanoes* 157–174 (Official Publ. Eur. Comm., 1998).
38. Gertisser, R., Preece, K. & Keller, J. The Plinian lower pumice 2 eruption, Santorini, Greece: magma evolution and volatile behaviour. *J. Volcanol. Geotherm. Res.* **186**, 387–406 (2009).
39. Lasaga, A. C. *Kinetic Theory in the Earth Sciences* (Princeton Univ. Press, 1998).
40. Zhang, Y. Diffusion in minerals and melts: theoretical background. *Rev. Mineral. Geochem.* **72**, 5–59 (2010).
41. Dohmen, R., Becker, H.-W., Meißner, E., Etzel, T. & Chakraborty, S. Production of silicate thin films using pulsed laser deposition (PLD) and applications to studies in mineral kinetics. *Eur. J. Mineral.* **14**, 1155–1168 (2002).
42. Cherniak, D. Cation diffusion in feldspars. *Rev. Mineral. Geochem.* **72**, 691–733 (2010).
43. Costa, F., Coogan, L. & Chakraborty, S. The time scales of magma mixing and mingling involving primitive melts and melt–mush interaction at mid-ocean ridges. *Contrib. Mineral. Petrol.* **159**, 371–387 (2010).
44. Costa, F. & Morgan, D. in *Timescales of Magmatic Processes: from Core to Atmosphere* (eds Dosseto, A., Turner, S. P. & Van Orman, J. A.) 125–159 (Wiley-Blackwell, 2010).

Serial translocation by means of circular intermediates underlies colour sidedness in cattle

Keith Durkin¹, Wouter Coppieeters¹, Cord Drögemüller², Naima Ahariz¹, Nadine Cambisano¹, Tom Druet¹, Corinne Fasquelle¹, Aynalem Haile³, Petr Horin⁴, Lusheng Huang⁵, Yohichiro Kamatani⁶, Latifa Karim¹, Mark Lathrop⁶, Simon Moser⁷, Kor Oldenbroek⁸, Stefan Rieder⁷, Arnaud Sartelet¹, Johann Sölkner⁹, Hans Stålhammar¹⁰, Diana Zelenika⁶, Zhiyan Zhang¹, Tosso Leeb², Michel Georges¹ & Carole Charlier¹

Colour sidedness is a dominantly inherited phenotype of cattle characterized by the polarization of pigmented sectors on the flanks, snout and ear tips¹. It is also referred to as 'lineback' or 'witrik' (which means white back), as colour-sided animals typically display a white band along their spine. Colour sidedness is documented at least since the Middle Ages and is presently segregating in several cattle breeds around the globe, including in Belgian blue and brown Swiss^{1,2}. Here we report that colour sidedness is determined by a first allele on chromosome 29 (*Cs₂₉*), which results from the translocation of a 492-kilobase chromosome 6 segment encompassing *KIT* to chromosome 29, and a second allele on chromosome 6 (*Cs₆*), derived from the first by repatriation of fused 575-kilobase chromosome 6 and 29 sequences to the *KIT* locus. We provide evidence that both translocation events involved circular intermediates. This is the first example, to our knowledge, of a phenotype determined by homologous yet non-syntenic alleles that result from a novel copy-number-variant-generating mechanism.

To gain insights into the molecular basis of colour sidedness (Fig. 1), we genotyped 21 colour-sided and 30 control Belgian blue animals with a custom-made 50K single nucleotide polymorphism (SNP) array³. As a result of segregation at the *roan* locus, Belgian blue animals are either black spotted (*r^{Bl}r^{Bl}*), blue spotted (*r^{Bl}r^{Wh}*) or white (*r^{Wh}r^{Wh}*)⁴. As white is epistatic to colour sidedness, we selected non-white control animals. We assumed autosomal dominant inheritance (*Cs* allele) and genetic homogeneity in Belgian blue, and thus scanned the genome of colour-sided animals for a shared haplotype (present in at least one copy) using the ASSDOM software (see Methods). This analysis yielded a single genome-wide significant signal ($P < 0.03$), mapping the *Cs* locus to bovine chromosome 29 (BTA29) (Fig. 2a and Supplementary Fig. 1). A sire transmitting the colour-sided phenotype to all its pigmented offspring was homozygous for the corresponding haplotype as expected. The shared haplotype spanned 1.9 Mb, and encompassed *LUZP2*—not known to be involved in pigmentation—as the only gene. Sequencing the *LUZP2* open reading frame (ORF) from colour-sided and control animals did not reveal any protein-sequence-altering variant (data not shown).

In an independent effort, we scanned the genome for copy number variants (CNVs) using a database of 50K SNP genotypes from >4,500 animals of different breeds and the PennCNV software⁵. Intriguingly, this analysis revealed a private duplication encompassing 26 SNPs on chromosome 6 (BTA6), shared by the 21 colour-sided Belgian blue animals (Fig. 2b). We confirmed and refined the boundaries of this ~480 kb duplication by comparative genome hybridization (CGH) of genomic DNA of a *Cs/Cs* and *+/+* Belgian blue animal

on a genome-wide Nimblegen tiling array (Fig. 2c). The corresponding CNV encompassed the *KIT* gene, known to be essential for melanocyte migration and survival⁶, and to be associated with coat colour variation^{7,8}.

To reconcile these apparently discrepant results, we performed fluorescence *in situ* hybridization (FISH) with bacterial artificial chromosome (BAC) clones mapping respectively to the BTA29 association interval (labelled with a red fluorophore), and to the BTA6 CNV interval (labelled with a green fluorophore) on lymphocytes of a *Cs/+* and a *+/+* animal. Overlapping red and green signals on one of the BTA29 homologues of the colour-sided animal demonstrated that the



Figure 1 | Colour-sided Belgian blue and brown Swiss animals. **a**, All represented Belgian blue animals are heterozygous *Cs/+*. In addition to the effect of the *Cs* locus, which causes polarization of pigmented sectors to the flanks, ears and snout, Belgian blue animals exhibit considerable polygenic variation in the degree of white spotting. The colour-sided phenotype of the two nearly completely white animals (extensive degree of white spotting) is apparent from the pigmentation of ear tips and snout. **b, c**, In brown Swiss, which are generally devoid of white spotting, homozygous *Cs/Cs* (**b**) and heterozygous *Cs/+* (**c**) animals differ by the extent of pigmentation, explaining why colour sidedness is also referred to as semi-dominant.

¹Unit of Animal Genomics, GIGA-R & Faculty of Veterinary Medicine, University of Liège (B34), 1 Avenue de l'Hôpital, 4000-Liège (Sart Tilman), Belgium. ²Institute of Genetics, University of Bern, Bremgartenstrasse 109a, Postfach 8466, 3001 Bern, Switzerland. ³International Center for Agricultural Research in the Dry Areas (ICARDA), PO Box 5466, Aleppo, Syria. ⁴Institute of Genetics, University of Veterinary and Pharmaceutical Sciences Brno, 612 42 Brno, Czech Republic. ⁵Key Laboratory for Animal Biotechnology of Jiangxi Province and the Ministry of Agriculture of China, NanChang, Jiangxi Province, 330045, China. ⁶Fondation Jean Dausset – Centre d'Etude du Polymorphisme Humain, 27 Rue Juliette Dodu, Paris 75012, France and CEA-IC Centre National de Génotypage, 2 Rue Gaston Crémieux, CP 5721, 91057 Evry Cedex, France. ⁷Swiss College of Agriculture, Bern University of Applied Sciences, 3052 Zollikofen, Switzerland. ⁸Centre for Genetic Resources the Netherlands (CGN), 6701 BH Wageningen, The Netherlands. ⁹Division of Livestock Sciences, University of Natural Resources and Applied Life Sciences, A-1180 Vienna, Austria. ¹⁰Viking Genetics, Örsnors, Box 64, S-532 21 Skara, Sweden.

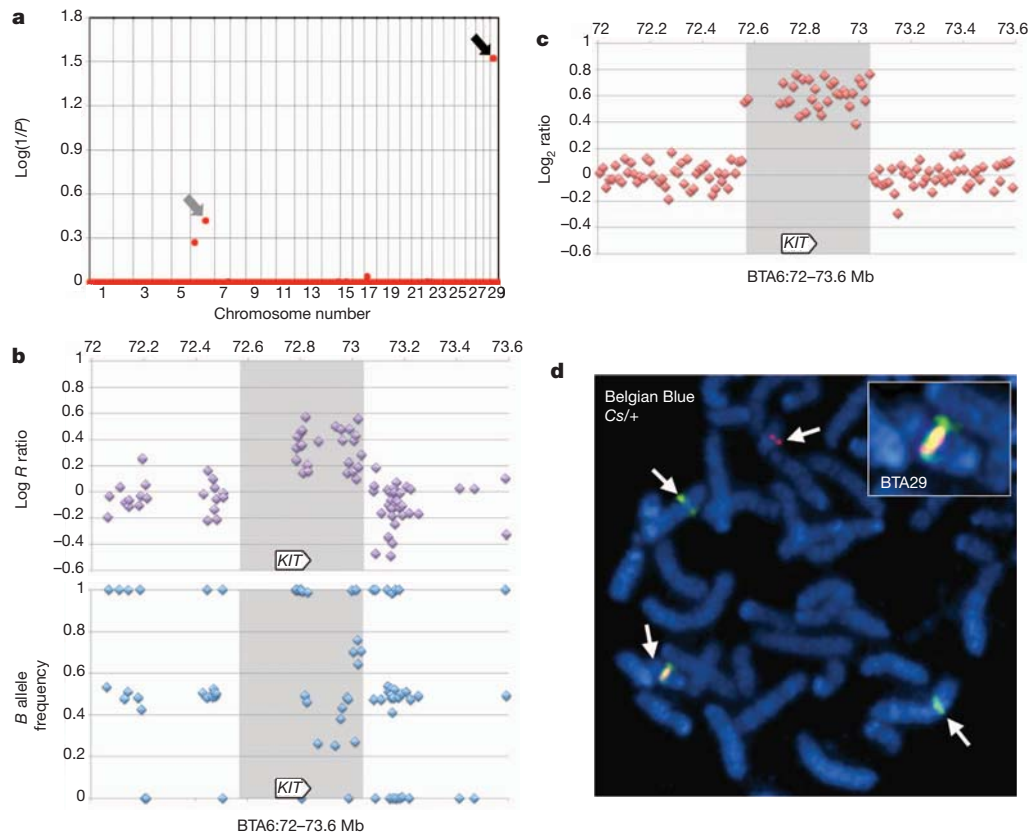


Figure 2 | Identification and mapping of the colour-sided locus in the Belgian blue. **a**, Genome-wide association mapping using ASSDOM (see Methods), revealing a genome-wide significant association ($P < 0.03$) on BTA29 (black arrow) and suggestive association on BTA6 (grey arrow). Vertical lines separate chromosomes. **b**, Detection of a duplication on the BTA6 segment shared by 21 colour-sided Belgian blue animals using 50K SNP array

CNV signal resulted from the translocation of the ~ 480 kb BTA6 segment to BTA29 (Fig. 2d). Retrospective examination of the SNP genotypes indicated that on average 3.17 of the 26 SNPs mapping to the ~ 480 kb duplication could not be called in colour-sided animals, yet the remaining genotypes were sufficient to yield a suggestive association signal (genome-wide $P = 0.37$), reflecting the sharing of the extra *KIT* haplotype by colour-sided animals (Fig. 2a).

To define the translocation breakpoints, we generated mate-pair libraries from self-ligated ~ 5 kb DNA fragments of a homozygous *Cs/Cs* Belgian blue animal, and generated ~ 10 Gb of sequence on an Illumina GAIIX instrument. We expected two clusters of aberrant mate pairs spanning the left and the right breakpoint resulting from the insertion of an intact BTA6 fragment in a BTA29 segment. We refer to the intact BTA6 fragment as $A-B-C-D-E$, the BTA29 segment in which it inserts as $\alpha-\beta$, and the resulting left and right breakpoints as $\alpha-A$ and $E-\beta$, respectively. However, we observed three aberrant mate-pair clusters corresponding respectively to $\alpha-D$, $E-A$ and $C-\beta$ fusions (Supplementary Fig. 2). The corresponding topology is most parsimoniously explained by assuming that: (1) the translocated BTA6 fragment circularized (generating fusion $E-A$); (2) reopened in the $C-D$ interval; and (3) integrated in the $\alpha-\beta$ BTA29 interval (generating fusions $\alpha-D$ and $C-\beta$; Fig. 3).

Using the genomic coordinates of the clustered mate pairs, we designed primer sets to amplify the three corresponding fusion points. Productive amplification was only achieved using genomic DNA from colour-sided Belgian blue animals, as expected (Supplementary Fig. 3A). We sequenced the corresponding PCR products to define the breakpoints at single-nucleotide resolution. The $E-A$ fusion was characterized by 2 bp micro-homology typical of non-homologous end joining (NHEJ)

data³ and PennCNV³. **c**, Confirmation and boundary definition using CGH on a Nimblegen bovine tiling array. Shaded areas in **b** and **c** mark the boundaries of the CNV as defined by CGH. **d**, Demonstration by FISH of the translocation of a *KIT*-encompassing BTA6 segment onto BTA29 in a heterozygous *Cs/+* Belgian blue animal. Magnification, $\times 100$.

(Supplementary Fig. 4A), whereas the $\alpha-D$ and $C-\beta$ fusions exhibited micro-duplications and micro-deletions reminiscent of replication-dependent microhomology-mediated break-induced replication (MMBIR) (Supplementary Fig. 4B)⁹. All breakpoints mapped to interspersed non-homologous repeat elements (data not shown).

The dominance of the *Cs* allele is expected to reflect a gain of function resulting from dysregulated expression of the translocated *KIT* gene. To verify the transcriptional competence of the translocated

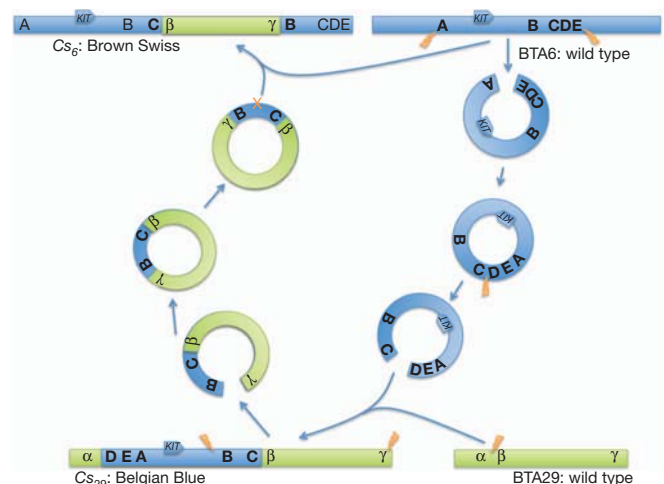


Figure 3 | Model for the generation of the colour-sided *Cs*₂₉ and *Cs*₆ alleles by serial translocation via circular shuttling intermediates.

KIT copy, we performed polymerase chain reaction with reverse transcription (RT-PCR) experiments using amplicons spanning two SNPs: the ss469414206 G-to-A transition in intron 1, and the ss469414207 C-to-T transition in intron 7. We extracted total RNA from pigmented and unpigmented skin sectors of a *Cs/+* colour-sided Belgian blue animal with the GG genotype on BTA6 and the A genotype on BTA29 ((GG)₆/A₂₉) for ss469414206 and the (CC)₆/T₂₉ genotype for ss469414207. The resulting pre-mRNA-dependent RT-PCR products were directly sequenced and the ratio of T/C and A/G species was estimated using Peakpicker¹⁰. T/A transcripts accounted for ~33% of the *KIT* output in both pigmented and unpigmented sectors, demonstrating the transcriptional potential of the translocated gene copy. Long-range RT-PCR analysis of the near complete *KIT* mRNA with primers located respectively in exon 1 and the 3' untranslated region (UTR) followed by amplicon sequencing did not show evidence for alternate transcripts in colour-sided animals (Supplementary Fig. 5).

Linkage analysis performed in a brown Swiss pedigree segregating for colour sidedness with microsatellite markers targeting candidate genes (*KIT*, *KITL*, *MITF*, *EDNRB*, *ADAMTS20*), yielded a log of odds (lod) score of 6.9 maximizing in the immediate vicinity of the *KIT* locus (Supplementary Table 1). We genotyped four colour-sided (three *Cs/Cs* and one *Cs/+*), and five control brown Swiss animals using a 50K SNP array³. Colour-sided *Cs/Cs* brown Swiss animals indeed shared a 2.0 Mb autozygous haplotype encompassing the *KIT* locus. At first glance, these findings suggested a distinct determinism of colour sidedness in Belgian blue and brown Swiss.

We analysed the corresponding SNP genotypes with PennCNV and performed CGH using genomic DNA from brown Swiss *Cs/Cs* versus *+/+* animals. This revealed the duplication of a ~120-kb BTA6 segment nested in the Belgian blue duplication, but excluding the *KIT*

gene. Intriguingly, it also revealed the duplication of a ~418-kb BTA29 segment immediately flanking the Belgian blue insertion site and encompassing the last four of twelve *LUZP2* exons (β - γ ; Fig. 4a, b). Moreover, fusion point C- β (but not α -C and D-A), specifying the Belgian blue *Cs* allele, could be amplified by PCR from genomic DNA of all examined brown Swiss colour-sided but not control animals (Supplementary Fig. 3). These findings established a clear link between the Belgian blue and brown Swiss *Cs* alleles.

We performed FISH analysis on lymphocytes of a *Cs/+* brown Swiss animal using BTA6 and BTA29 BAC clones. Remarkably, we observed overlapping red and green signals on one of the BTA6 homologues, hence revealing the translocation of a BTA29 fragment on BTA6 (Fig. 4c). We generated mate-pair libraries from self-ligated ~5-kb and ~2-kb DNA fragments of a brown Swiss *Cs/Cs* animal and generated ~15 Gb of sequence on an Illumina GAIIx instrument (Supplementary Fig. 2). Analysis of the resulting sequence traces revealed two aberrant mate-pair clusters. The first corresponds to the Belgian blue C- β fusion point, previously detected by PCR. The second corresponds to a novel B- γ fusion point (Supplementary Fig. 2). The most parsimonious explanation accounting for all the data is that the brown Swiss *Cs* allele derives from the Belgian blue *Cs* allele by (1) excision of the B-C- β - γ fragment from the Belgian blue *Cs* allele on BTA29, (2) circularization, and (3) re-integration in the wild-type *KIT* locus by homologous recombination. This would result in a novel *Cs* allele mapping to the *KIT* locus and characterized by tandem duplicates of the B-C segment flanking the translocated β - γ BTA29 fragment (A-B-C- β - γ -B-C-D-E; Fig. 3).

Our model predicts that the C and B BTA6 segments immediately flanking β - γ carry the same haplotype as the *Cs*-allele of Belgian blue. To test this, we developed long-range PCR assays that would specifically

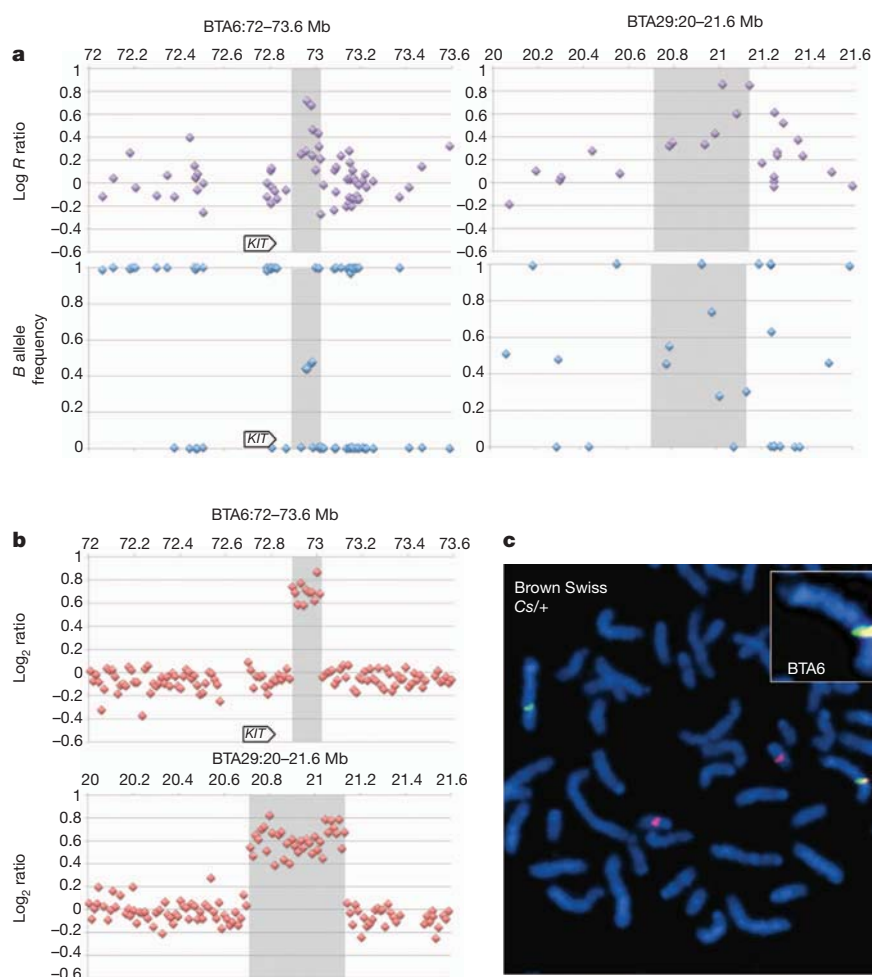


Figure 4 | Identification and mapping of the colour-sided locus in brown Swiss animals.

a, PennCNV evidence for BTA6 and BTA29 duplications in brown Swiss colour-sided animals. The slightly different distribution of BTA6 SNPs when compared to Fig. 2 is due to the use of a different version of the 50K SNP array¹.

b, Confirmation and boundary definition of the *Cs*-associated BTA6 and BTA29 duplication in brown Swiss animals using CGH on Nimblegen bovine tiling arrays. The shaded areas in **a** and **b** mark the boundaries of the CNV, as defined by CGH.

c, Demonstration by FISH of the translocation of a BTA29 segment onto BTA6 in a heterozygous *Cs/+* brown Swiss animal.

amplify B, C, B and C segments from genomic DNA of a Cs/Cs brown Swiss animal, and—by sequencing the corresponding amplicons—determined the genotype of B and B and C and C for four SNP positions heterozygous in the high-throughput sequence reads of the brown Swiss Cs/Cs animal. We then determined the genotype of the Belgian blue Cs allele for the corresponding variants and showed that it matched the C and B segments of the brown Swiss Cs allele, in agreement with our model of homologous-recombination-dependent resolution of the circular intermediate (Fig. 3 and Supplementary Fig. 6).

Using the mate-pair genomic coordinates, we designed primer pairs to amplify the γ -B fusion point. As expected, productive amplification was only achieved using genomic DNA from colour-sided brown Swiss animals (Supplementary Fig. 3). We sequenced the corresponding PCR products to define the breakpoints at single-nucleotide resolution. The γ -B fusion presented hallmarks typical of microhomology mediated break induced replication (MMBIR) (Supplementary Fig. 4).

We obtained genomic DNA from colour-sided animals from seven additional cattle breeds and domestic yaks (Supplementary Fig. 7), which we tested by PCR for the presence of the two Belgian-blue-specific fusion points (E-A, α -D), the brown-Swiss-specific fusion point (γ -B), and the Belgian blue/brown Swiss shared fusion point (C- β). Colour-sided Dutch witrik and Ethiopian fogera animals were shown to carry the Belgian blue Cs allele (Cs_{29}), Austrian pustertaler sprinzen, Czech red-spotted cattle and French vosgienne the brown Swiss Cs allele (Cs_6), and Irish moiled, Swedish mountain and domestic yak carried both the Cs_{29} and Cs_6 alleles (Supplementary Fig. 3). We assume that Cs_{29} and Cs_6 alleles were introgressed in yak after domestication via well-documented hybridization of *Bos taurus* and *Bos grunniens*. These findings indicate that the Cs_{29} and Cs_6 alleles account for most if not all colour sidedness in cattle.

Analysis of colour sidedness has revealed a novel CNV-generating translocation mechanism involving circular intermediates. Whether this is a bovine idiosyncrasy or a more common mechanism remains to be determined. That some CNVs reflect translocation events is well established. As an example, 75 probable dispersed duplications in the human genome have been reported¹¹, as well as at least four interchromosomal duplications^{12,13}. We ourselves performed genome-wide association mapping between the SNP and CNV genotype in human and cattle, and observed 21 and 4 putative 'trans-associations' (that is, non-syntenic CNV-defining SNPs and associated SNPs), respectively (Supplementary Material). Some of these dispersed duplications might involve circular intermediates. In support of the occurrence of other instances involving circular intermediates are recent observations of repeated translocation of five clustered ORFs in wine and bioethanol strains of *Saccharomyces cerevisiae*, apparently via resolution of circular intermediates¹⁴. The same mechanism may contribute to somatic mutations in tumours. Indeed, episomes with the *NUP214-ABL1* fusion gene (observed in ~6% of T-cell acute lymphoblastic leukaemia) have been proposed to result from circularization and excision of a chromosome 9 segment bounded by the *NUP214* and *ABL1* genes, and to reintegrate ectopically by the same resolving mechanism proposed for colour sidedness in at least some patients¹⁵. The repatriation of exogenous sequences (including exons) back to the original chromosomal location via circular 'shuttling' intermediates suggests that this mechanism might underlie a specific mode of exon shuffling.

METHODS SUMMARY

A custom 50K SNP array³ was used to genotype 21 colour-sided and 30 control Belgian blue animals. The genome of colour-sided animals was scanned for a shared haplotype using the ASSDOM software. In the brown Swiss, microsatellites adjacent to five candidate genes were genotyped in three half-sibling families and two point linkage calculated using the Merlin software¹⁶. CNVs were identified in the 50K SNP array data using the PennCNV software⁵. Array CGH was carried out on a custom 2.1 M oligonucleotide array (Roche-Nimblegen) with a non-colour-sided Belgian blue used as the reference in each hybridization. Metaphase spreads

were generated from short-term lymphocyte cultures. BACs from the duplicated regions on BTA6 and BTA29 were identified using end sequences from the bovine RPCI42 BAC library. BACs were labelled with the appropriate fluorochrome by nick translation (Abbott Molecular), hybridized to the metaphase spreads and examined by fluorescent microscopy. Mate-pair libraries were generated using the Illumina mate-pair library kit v.2 for a Cs/Cs Belgian blue and a Cs/Cs brown Swiss animal. A paired-end library was also generated for a Cs/+ Belgian blue using the Illumina paired-end kit. Sequencing was carried out on an Illumina GAIIx instrument. PCR products spanning regions of interest were purified using QIAquick PCR purification kit (Qiagen), sequenced using Big Dye terminator cycle-sequencing kit v.3.1 (Applied Biosystems) and run on an ABI PRISM 3730 DNA analyser (Applied Biosystems). RNA was extracted from skin using the RNeasy fibrous tissue mini kit (Qiagen) and cDNA was synthesized using SuperScript II first-strand Synthesis SuperMix (Invitrogen). Long-range PCR was carried out using the Expand Long Template PCR System (Roche).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 September; accepted 5 December 2011.

- Olson, T. *The Genetics of Cattle* (CAB International, 1999).
- Porter, V. & Mason, I. L. *Mason's World Dictionary of Livestock Breeds, Types, and Varieties* (CAB International, 2002).
- Charlier, C. *et al.* Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genet.* **40**, 449–454 (2008).
- Charlier, C. *et al.* Microsatellite mapping of the bovine roan locus: a major determinant of white heifer disease. *Mamm. Genome* **7**, 138–142 (1996).
- Wang, K. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
- Yoshida, H. *et al.* Melanocyte migration and survival controlled by *SCF/c-kit* expression. *J. Invest. Dermatol. Symp. Proc.* **6**, 1–5 (2001).
- Brooks, S. A. *et al.* A chromosome inversion near the *KIT* gene and the Tobiano spotting pattern in horses. *Cytogenet. Genome Res.* **119**, 225–230 (2007).
- Marklund, S. *et al.* Molecular basis for the dominant white phenotype in the domestic pig. *Genome Res.* **8**, 826–833 (1998).
- Hastings, P. J., Lupski, J. R., Rosenberg, S. M. & Ira, G. Mechanisms of change in gene copy number. *Nature Rev. Genet.* **10**, 551–564 (2009).
- Ge, B. *et al.* Survey of allelic expression using EST mining. *Genome Res.* **15**, 1584–1591 (2005).
- Conrad, D. F. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704–712 (2010).
- Liu, G. E. *et al.* Analysis of copy number variations among diverse cattle breeds. *Genome Res.* **20**, 693–703 (2010).
- Liu, G. E. *et al.* Analysis of recent segmental duplications in the bovine genome. *BMC Genomics* **10**, 571 (2009).
- Borneman, A. R. *et al.* Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet.* **7**, e1001287 (2011).
- Graux, C. *et al.* Heterogeneous patterns of amplification of the *NUP214-ABL1* fusion gene in T-cell acute lymphoblastic. *Leukemia* **23**, 125–133 (2009).
- Abecasis, G. *et al.* Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet.* **30**, 97–101 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by the Walloon Direction General Operationnelle Agriculture, Ressources Naturelles et Environnement (grants D31-1206 and D31-1214) and the Swiss National Science Foundation (grant 31003A_133034 to T.L.). We are grateful for the support of the GIGA-R Genomics platform, to P. Melvin, to G. Mekuriaw and staff of the Andassa livestock research center, Viking Genetics, Swissgenetics and the Schweizer Braunviehzuchtverband for their assistance in sample collection, to J. Vermeesch and B. Nowakowska for their assistance with array CGH.

Author Contributions K.D. performed the PennCNV, CGH, FISH, next-generation sequencing, breakpoint, DNA sequence variant and expression analyses. W.C. performed the bioinformatic analyses of the high-throughput sequencing reads. A.H., P.H., L.H., K.O., A.S., J.S. and H.S. provided samples from colour-sided animals. C.D., S.M. and S.R. performed linkage analyses in brown Swiss animals. T.L. supervised the work in Switzerland. C.C., M.G., T.D. and Z.Z. performed the association analyses. N.A. and N.C. performed SNP genotyping. C.F. and L.K. performed amplicon sequencing. Y.K., M.L. and D.Z. provided human CNV genotype data. K.D., M.G. and C.C. analysed data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.G. (michel.georges@ulg.ac.be).

METHODS

Association and linkage mapping of the colour-sided locus. SNP genotyping was conducted using custom-made 50K Infinium SNP arrays³ used according to the instructions of the manufacturer. Microsatellite genotyping was conducted as previously described¹⁷. Association analysis was conducted using ASSDOM. ASSDOM searches for chromosome segments devoid of SNPs for which cases have alternate homozygous genotype (say 11 versus 22), excluding the sharing of an identical-by-descent (single-copy) haplotype. Intervals bounded by such excluding SNPs receive a score corresponding to $\sum_{i=1}^k \log(1-p_i^n)$, where p_i is the frequency of the allele missing among n cases, estimated in m controls. The genome-wide statistical significance of the 'non-exclusion' signal is determined by phenotype permutation of the disease status between the n cases and m controls. Two-point linkage analyses were conducted with Merlin¹⁶.

Prediction of CNVs from SNP genotype data. The log R ratio signal intensity and B allele frequency from a custom 50K SNP array³ were obtained using Illumina BeadStudio software. PennCNV, a hidden Markov model based approach that takes into account signal intensity, allelic intensity ratio, distance between markers and allele frequency⁵ was used to call CNVs. Regions of the genome that showed evidence of copy number change were inspected in greater detail. Plots of log R ratio and B allele frequency were examined in BeadStudio (Illumina) and the region checked in the University of California, Santa Cruz (UCSC) genome browser (<http://genome.ucsc.edu/>).

Detection of CNVs by CGH. Array CGH was carried out on a custom 2.1 M oligonucleotide array (Roche-Nimblegen) based on the UMD 3.0 bovine assembly. The array contained 2,152,422 probes (50–75 mers) with a median spacing of 1,160 bp. The reference animal used in hybridizations was a non-colour-sided Belgian blue. Genomic DNA labelling (Cy3 for sample and Cy5 reference), hybridization and washing were performed according to the manufacturer's instructions and have been described elsewhere¹⁸. Slides were scanned using a GenePix 4000B 5 μ m microarray scanner (Axon Instruments). Images were processed using NimbleScan software (Roche-Nimblegen). Spatial correction was applied and data normalized¹⁹, segmentation was performed using the DNACopy algorithm²⁰. The log₂ ratios for each oligonucleotide were also examined visually for evidence of a change in copy number in regions of interest.

FISH. Peripheral blood was obtained from colour-sided and wild-type Belgian blue and brown Swiss animals. Pokeweed-stimulated lymphocyte cultures were established and chromosome spreads prepared for colour-sided and wild-type animals in each breed following standard cytogenetic procedures. End sequences from the RPCI42 bovine BAC library and the duplicated regions on BTA6 and BTA29 in the colour-sided Belgian blue and brown Swiss were downloaded from the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>). BLAST was used to identify end sequences located in the duplicated regions. The BACs 160M9 and 156I13 overlap and cover the region chr6:72,566,605–72,817,995 (bosTau4), while the BACs 37P11 and 116G8 also overlap and cover the region chr29:20,772,406–21,035,251 (bosTau4). BAC clones were initially cultured at 37 °C in 1 ml of 2YT media containing 30 μ l ml⁻¹ chloramphenicol. The cultures were plated on lysogeny broth (LB) agar plates with 30 μ l ml⁻¹ chloramphenicol to obtain single colonies, the identity of the BACs was confirmed by PCR using primers designed within the area encompassed by the respective BAC (Supplementary Table 2). The positive clones were then used to inoculate 100 ml of 2YT media with 30 μ l ml⁻¹ chloramphenicol. Following 24 h at 37 °C with constant agitation, DNA was extracted using the Qiagen midiprep kit, following the manufacturer's instructions. The DNA was labelled with the nick translation kit from Abbott Molecular, using the manufacturer's protocol. DNA from BACs 160M9 and 156I13 was mixed and labelled with spectrum green (Abbott Molecular) while 37P11 and 116G8 were mixed and labelled with spectrum orange (Abbott Molecular). Labelled DNA (100 ng) was combined with 1 μ g of bovine Cot-1 DNA and 2 μ g of bovine genomic DNA, precipitated then resuspended in 3 μ l of purified water and 7 μ l of hybridization buffer (Abbott Molecular). The separate probes were denatured at 73 °C for 5 min and then combined on a slide containing metaphase spreads. These slides had been denatured in 70% formamide in 2 \times SSC at 75 °C for 5 min followed by dehydration in ethanol. A coverslip was secured with rubber cement and the slide incubated overnight at 37 °C in a humidified chamber. Slides were then washed in 0.4 \times SSC/0.3% Tween-20 at 73 °C for 1–3 s followed by washing in 2 \times SSC/0.1% Tween-20 at room temperature (~18 °C) for 1–3 s and air dried. The slides were counterstained with DAPI II (Abbott Molecular) and visualized by fluorescent microscopy.

Next-generation sequencing of mate-pair and paired-end libraries. Mate-pair libraries with different insert sizes were generated using the Illumina mate-pair library kit, v.2. The manufacturer's instructions were followed except for the step involving fragmentation of the circularized DNA, for which a bioruptor sonicator

UCD-200 (Diagenode) was used. The 300 μ l sample was placed in a 1.5 ml Eppendorf tube and sonicated for 8 min with the instrument set to high and a cycle of 30 s on and 30 s off. A 5-kb insert library was generated for a *Cs/Cs* Belgian blue animal and a 2-kb and 5-kb for a *Cs/Cs* brown Swiss animal. A paired-end library with a 400 bp insert size was also generated for a *Cs/+* Belgian blue animal using the Illumina paired-end kit, following the instructions of the manufacturer. The resulting libraries were quantified using Pico-Green (Quant-it, Invitrogen) and the Agilent 2100 Bioanalyzer High Sensitivity DNA kit (Agilent Technologies). Sequencing was carried out on an Illumina GAIIX instrument. Mapping of the 36 bp from each end of the mate-pair libraries and the 110 bp from the ends of the paired-end library was performed using the BWA tool²¹. Breakpoints were identified by visually inspecting the mate pairs using the integrative genomics viewer²² in the ~2-Mb region surrounding the *Cs*-specific duplications in Belgian blue and brown Swiss animals and looking for discordant mate pairs.

PCR amplification of translocation breakpoints. PCR primers were designed to span each of the breakpoints identified by mate-pair sequencing. The primers were tested on genomic DNA from colour-sided and wild-type animals. PCR products were visualized on a 2% agarose gel. Primers with amplification confined to colour-sided animals were purified using the QIAquick PCR purification kit (Qiagen), where multiple bands were observed the relevant band was excised and purified using the QIAquick gel extraction kit (Qiagen). The fragments were then sequenced using Big Dye terminator cycle-sequencing kit v.3.1 (Applied Biosystems) with the purified reaction run on a ABI PRISM 3730 DNA analyser (Applied Biosystems). Primers used are listed in Supplementary Table 2.

Analysis of *Cs*₂₉-derived *KIT* transcripts. To ensure that no mutations were present in the coding sequence of the *Cs*₂₉-specific *KIT* gene, primers were designed to amplify all the exons and the 3' UTR from genomic DNA (Supplementary Table 2). The resulting PCR products were sequenced as outlined above, which did not reveal any protein altering DNA sequence variant. To examine expression of *KIT* from the *Cs*₂₉ allele, a small biopsy of skin was removed from the back (white skin) and side (pigmented skin) of a colour-sided Belgian blue animal (the relevant ethical procedures were adhered to). The samples were immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction. The tissue was homogenized using a Tissue Lyser (Qiagen), total RNA was extracted using the RNeasy fibrous tissue mini kit (Qiagen), following the manufacturer's instructions. First-strand cDNA was synthesized using SuperScript II first-strand Synthesis SuperMix (Invitrogen) with a mixture of random hexamers and oligo(dT) primers. Genomic DNA was extracted from whole blood from the same animal using standard phenol-chloroform extraction. The intronic regions of the *KIT* gene were searched for SNPs using the mate-pair sequences produced from the *Cs/Cs* Belgian blue animal. Primers were designed around suitable SNPs and sequenced in the tissue donor. To establish the genotype of the *Cs*₂₉ allele, multiple colour-sided animals were sequenced. For SNP ss469414207 (C/T) 40 colour-sided animals were sequenced. For all *Cs/Cs* animals, the C and T alleles had equal peak height, while in the remaining *Cs/+* animals the T allele produced the smaller peak. Additionally seven non colour-sided Belgian blue animals were also sequenced and none possessed the T allele. For SNP ss469414206 (G/A), two *Cs/Cs* and three *Cs/+* animals were sequenced and showed a pattern of peak heights consistent with the *Cs*₂₉ allele having the A genotype. To determine if the *Cs*₂₉ allele was expressed, the cDNA from the skin biopsies was amplified using primers spanning both SNPs and sequenced as outlined above. Moreover, and to ensure the integrity of *Cs*₂₉-derived *KIT* transcripts, primers located in the first exon and the 3' UTR (Supplementary Table 2) were used with the Expand Long Template PCR System (Roche) to amplify the near full-length *KIT* cDNA. The product was run on 1% gel and examined for evidence of alternative splicing.

Genotyping the duplicated B and C segments of the *Cs*₆ allele. Long-range PCR was carried out using the Expand Long Template PCR System (Roche). The genomic DNA used as template was extracted using QIAamp DNA Mini columns (Qiagen) following the manufacturer's instructions to produce high molecular weight DNA. For each reaction the following mix was prepared: 2 μ l Buffer 1, 140 μ M dNTP, 120 nM upstream and downstream primer, 0.3 μ l enzyme mix and 100 ng genomic DNA, final volume was 20 μ l. Extension time and the thermal profile recommended by the manufacturer was followed. Following 30 cycles 10 μ l of the product was run on a 0.8% agarose gel to check for amplification, the remaining reaction mix was retained. The PCR primers spanning the relevant SNP were then used in a conventional PCR reaction using between 0.1 and 0.5 μ l (amount determined by intensity of band) of the long-range product as template. When long-range PCR product was used as template the number of cycles was reduced to 20. Clean-up and sequencing was carried out as outlined above. Used primers are reported in Supplementary Table 2.

17. Drögemüller, C. *et al.* Genetic mapping of the belt pattern in Brown Swiss cattle to BTA3. *Anim. Genet.* **40**, 225–229 (2009).

18. Selzer, R. *et al.* Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosom. Cancer* **44**, 305–319 (2005).
19. Workman, C. *et al.* A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol.* **3**, 0048 (2002).
20. Olshen, A. *et al.* Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5**, 557–572 (2004).
21. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
22. James, T. *et al.* Integrative genomics viewer. *Nature Biotechnol.* **29**, 24–26 (2011).

Neuron-type-specific signals for reward and punishment in the ventral tegmental area

Jeremiah Y. Cohen^{1*}, Sebastian Haesler^{1*}, Linh Vong², Bradford B. Lowell² & Naoshige Uchida¹

Dopamine has a central role in motivation and reward. Dopaminergic neurons in the ventral tegmental area (VTA) signal the discrepancy between expected and actual rewards (that is, reward prediction error)^{1–3}, but how they compute such signals is unknown. We recorded the activity of VTA neurons while mice associated different odour cues with appetitive and aversive outcomes. We found three types of neuron based on responses to odours and outcomes: approximately half of the neurons (type I, 52%) showed phasic excitation after reward-predicting odours and rewards in a manner consistent with reward prediction error coding; the other half of neurons showed persistent activity during the delay between odour and outcome that was modulated positively (type II, 31%) or negatively (type III, 18%) by the value of outcomes. Whereas the activity of type I neurons was sensitive to actual outcomes (that is, when the reward was delivered as expected compared to when it was unexpectedly omitted), the activity of type II and type III neurons was determined predominantly by reward-predicting odours. We ‘tagged’ dopaminergic and GABAergic neurons with the light-sensitive protein channelrhodopsin-2 and identified them based on their responses to optical stimulation while recording. All identified dopaminergic neurons were of type I and all GABAergic neurons were of type II. These results show that VTA GABAergic neurons signal expected reward, a key variable for dopaminergic neurons to calculate reward prediction error.

Dopaminergic neurons fire phasically (100–500 ms) after unpredicted rewards or cues that predict reward^{1–3}. Their response to reward is reduced when a reward is fully predicted. Furthermore, their activity is suppressed when a predicted reward is omitted. From these observations, previous studies hypothesized that dopaminergic neurons signal discrepancies between expected and actual rewards (that is, they compute reward prediction error (RPE)), but how dopaminergic neurons compute RPE is unknown.

Dopaminergic neurons make up about 55–65% of VTA neurons; the rest are mostly GABAergic inhibitory neurons^{4–6}. Many addictive drugs inhibit VTA GABAergic neurons, which increases dopamine release (called disinhibition), a potential mechanism for reinforcing the effects of these drugs^{7–12}. Despite the known role of VTA GABAergic neurons inhibiting dopaminergic neurons *in vitro*¹³, little is known about their role in normal reward processing. One obstacle has been the difficulty of identifying different neuron types with extracellular recording techniques. Conventionally, spike waveforms and other firing properties have been used to identify presumed dopaminergic and GABAergic neurons^{1,2,14,15}, but this approach has been questioned recently^{5,16}. We thus aimed to observe how dopaminergic and GABAergic neurons process information about rewards and punishments.

We classically conditioned mice with different odour cues that predicted appetitive or aversive outcomes. The possible outcomes were big reward, small reward, nothing, or punishment (a puff of air delivered to the animal’s face). Each behavioural trial began with a conditioned stimulus (CS; an odour, 1 s), followed by a 1-s delay and

an unconditioned stimulus (US; the outcome). Within the first two behavioural sessions, mice began licking towards the water-delivery tube in the delay before rewards arrived, indicating that they quickly learned the CS–US associations (Fig. 1). The lick rate was significantly higher preceding big rewards than small ones (paired *t*-tests between lick rates for big versus small rewards for each session, $P < 0.05$ for each mouse).

We recorded the activity of VTA neurons while mice performed the conditioning task. All 95 neurons showed task-related responses (analysis of variance (ANOVA), all $P < 0.001$), thus all recorded neurons were used in the following analyses. Observing the temporal profiles of responses in trials with rewards, we found neurons that showed firing patterns that resemble those of dopaminergic neurons found in non-human primates^{1,2,15}. These neurons were excited phasically by reward-predicting stimuli or reward (Fig. 2a, top). We also found many neurons with firing patterns distinct from typical dopaminergic neurons. These neurons showed persistent excitation during the delay before rewards, in response to reward-predicting odours (Fig. 2a, middle). Other neurons showed persistent inhibition to reward-predicting odours (Fig. 2a, bottom). To characterize the responses of the population, we measured the temporal response profile of each neuron during big-reward trials by quantifying firing rate changes from baseline in 100-ms bins using a receiver operating characteristic (ROC) analysis (Fig. 2b and Supplementary Fig. 1). We calculated the area under the ROC curve (auROC) at each time bin. Values greater than 0.5 indicate increases in firing rate relative to baseline, whereas values less than 0.5 indicate decreases.

To classify these response profiles, we used principal component analysis (PCA) followed by unsupervised, hierarchical clustering. This yielded three clusters of neurons that were separated according to (1) the magnitude of activity during the delay between CS and US, and (2) the magnitude of responses to the CS or US (Fig. 2c). Forty-nine neurons (52%) were classified as type I, which showed phasic responses. Twenty-nine neurons (31%) were classified as type II, which showed sustained excitation to reward-predicting odours, whereas 17 neurons (18%) were classified as type III, which showed sustained inhibition (Fig. 2d).

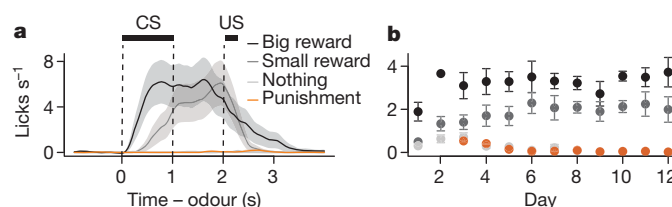


Figure 1 | Odour-outcome association task in mice. **a**, Licking behaviour from a representative experimental session. Black bars indicate CS and US delivery. Shaded regions around lick traces denote standard error of the mean (s.e.m.). **b**, Mean \pm s.e.m. licks during the delay between CS and US as a function of days of the experiment across animals.

¹Department of Molecular and Cellular Biology, Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA. ²Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA.

*These authors contributed equally to this work.

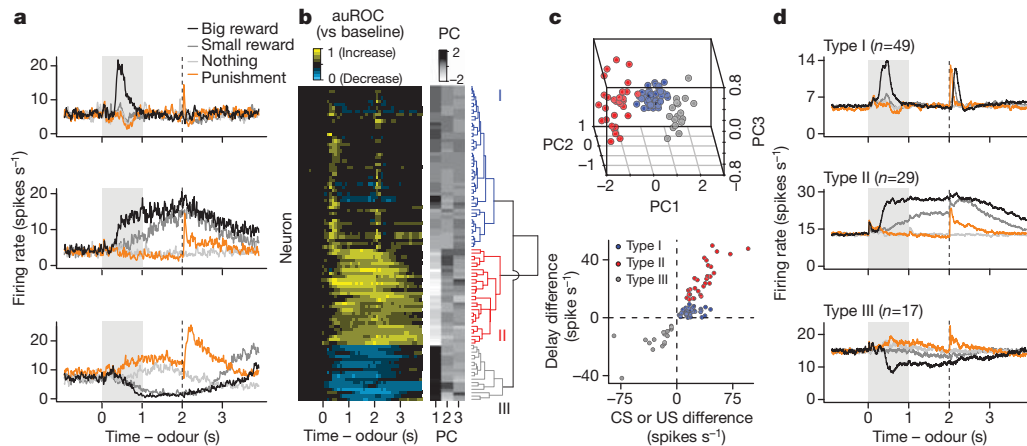
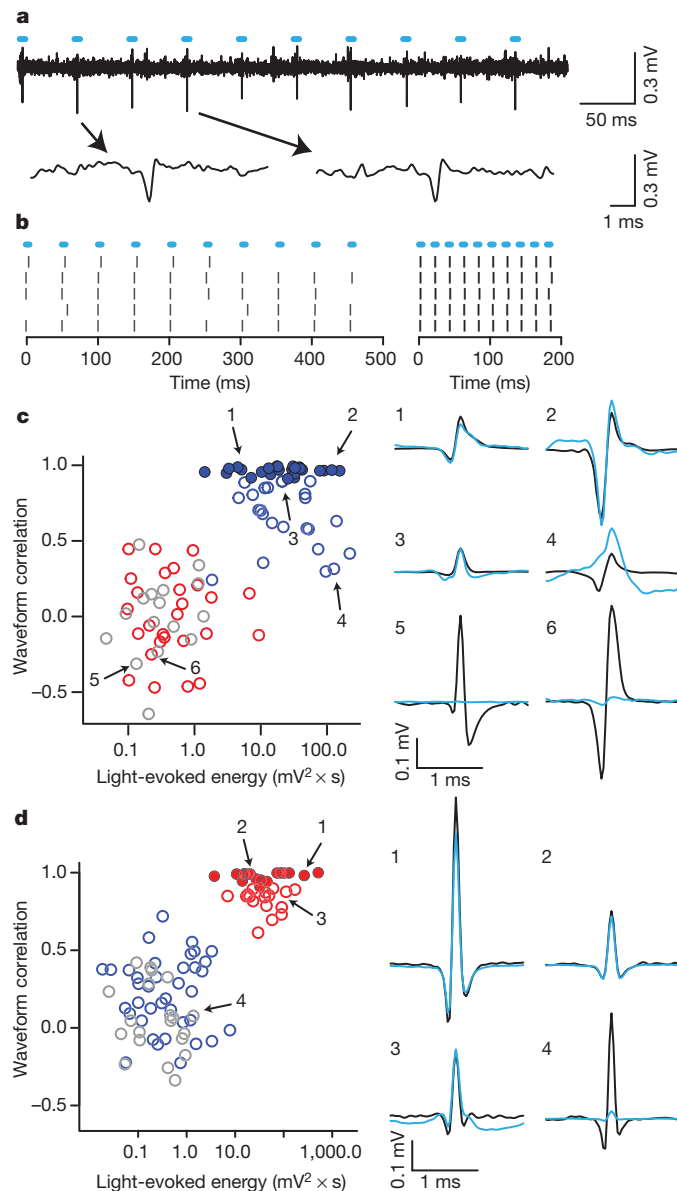


Figure 2 | VTA neurons show three distinct response types. **a**, Responses of example neurons. **b**, Responses of all neurons. Yellow, increase from baseline; cyan, decrease from baseline. Each row represents one neuron. The similarity order of the three main clusters is arranged to match the order presented in **a**. **c**, Top: the first three principal components of the auROC curves. Points are

coloured based on hierarchical clustering from the dendrogram. Bottom: classification of neurons based on response differences between big-reward and no-reward trials during the delay versus during the CS or US. **d**, Average firing rates from type I–III neurons.



To identify dopaminergic neurons, we expressed channelrhodopsin-2 (ChR2), a light-gated cation channel^{17,18}, in dopaminergic neurons (see Methods). We confined expression to dopaminergic neurons by injecting adeno-associated virus containing FLEX-ChR2 (AAV-FLEX-ChR2)¹⁹ into transgenic mice expressing Cre recombinase under the control of the promoter of the dopamine transporter (*DAT*; also called *Slc6a3*) gene (Supplementary Figs 2 and 3). For each neuron, we measured the response to light pulses and the shape of spontaneous spikes. We observed many neurons that fired after light pulses (Fig. 3a, b). We calculated the correlation between the spontaneous spike waveform and light-evoked voltage response and plotted it against the energy of light-evoked responses for each recording (Fig. 3c). This yielded two distinct clusters: one that showed significant responses to light pulses and one that did not. To identify dopaminergic neurons stringently, we applied the criterion that the light-evoked waveform must look almost identical to the spontaneous waveform (correlation coefficient >0.9). Twenty-six neurons met this criterion (filled blue points in Fig. 3c). Consistent with direct light activation rather than indirect, synaptic activation, all 26 neurons showed light-evoked spikes within a few milliseconds of light onset with small jitter, and followed high-frequency light stimulation of 50 Hz (Supplementary Fig. 4). These properties strongly indicate that these 26 neurons expressed ChR2. We therefore designate these 26 neurons as identified dopaminergic neurons. All identified dopaminergic neurons were of type I. Conversely, none of type II or III neurons was activated by light (red and grey points in Fig. 3c).

Next, we asked whether GABAergic neurons could be mapped to type II or III neurons. We recorded from 92 VTA neurons in mice expressing Cre recombinase under the control of the endogenous vesicular γ -aminobutyric acid (GABA) transporter (*Vgat*; also called *Slc32a1*) gene. These mice showed similar licking behaviour to *DAT*-Cre mice (Supplementary Fig. 5). We applied the PCA parameters

Figure 3 | Identifying dopaminergic and GABAergic neurons. **a**, Voltage trace from 10 pulses of 20 Hz light stimulation (cyan bars). Two light-triggered spikes are shown below. **b**, Response from this neuron to 20 Hz (left) and 50 Hz (right) stimulation. Ticks represent spikes. **c**, Quantification of light-evoked responses and mapping of response types in *DAT*-Cre mice. Blue, type I; red, type II; grey, type III neurons. Identified dopaminergic neurons are indicated by filled circles. Abscissa: energy (integral of the squared voltage values, $\int v^2 dt$) of the light-evoked response from each neuron. Ordinate: cross-correlation between the mean spontaneous spike and the light-evoked response. Example neurons are shown to the right (black, spontaneous spikes; cyan, light-evoked voltages). **d**, Light-evoked responses in *Vgat*-Cre mice. Conventions are the same as in **c**.

from the 95 neurons from *DAT-Cre* mice to the 92 neurons from *Vgat-Cre* mice. This yielded 38 type I neurons, 34 type II neurons and 20 type III neurons. Using the same criteria for GABAergic neurons as we used for dopaminergic neurons, we identified 17 GABAergic neurons (Fig. 3d and Supplementary Fig. 4). All 34 type II neurons fell in the upper cluster in Fig. 3d. We also found type I neurons that were inhibited by optical stimulation, consistent with local GABAergic stimulation (Supplementary Fig. 6).

Our data set of identified dopaminergic neurons allows us to characterize their diversity. We observed that some were excited by reward, some were excited by a reward-predicting CS, and some were excited by both (Fig. 4a–c). Although previous studies in non-human primates found similar variability^{20,21} (Supplementary Fig. 7), this result may suggest that some dopaminergic neurons do not strictly follow canonical RPE coding. However, the US responses may be due to the delay between CS and US, known to increase the US response due to temporal uncertainty²⁰. In addition, this diversity was correlated with the effect of training that occurred over several days across the population of dopaminergic neurons, even after animals had reached asymptotic behavioural performance (Fig. 1b). Soon after reaching a behavioural performance criterion, many dopaminergic neurons showed stronger responses to US over CS but the preference gradually shifted to CS over several days (Fig. 4d; Pearson correlation, $r = 0.42$, $P < 0.05$). This is consistent with a previous study in non-human primates that showed US responses gradually disappear over >1 month of training²¹. Thus, identified dopaminergic neurons appear to respond to CS and US similarly to those reported in non-human primate studies.

Another important response property that supports RPE coding in dopaminergic neurons is their decrease in firing rate when an expected reward is omitted^{1,3}. We thus omitted reward unexpectedly on 10% of big-reward trials in some sessions. Fifteen of seventeen dopaminergic neurons showed a decrease in firing rate upon reward omission relative to reward delivery (Fig. 4f, g). The two dopaminergic neurons that

were not modulated by reward omission were excited by big-reward CS, but fired close to 0 spikes s^{-1} otherwise; the low firing rate at the time of reward left little room to ‘dip’ further. We obtained similar results when we compared the firing rate upon reward omission to the baseline firing rate (9 of 17 neurons $P < 0.05$, t -test; mean auROC = 0.407, $t_{16} = 2.56$, $P < 0.05$; Supplementary Fig. 8a, b). Thus, most dopaminergic neurons coded RPE when expected reward was omitted.

GABAergic neurons showed persistent activity during the delay period, which parametrically encoded the value of upcoming outcomes (paired t -tests between no-, small- and big-reward trials, all $P < 0.001$ for 16 of 17 identified GABAergic neurons, Supplementary Fig. 7a; regression slopes, Supplementary Fig. 10i). This suggests that these neurons encode expectation about rewards. If this is the case, one prediction is that the activity of these neurons is not modulated by delivery or omission of reward. Indeed, GABAergic (and unidentified type II) and type III neurons were not significantly modulated by the presence or absence of reward itself (Fig. 4f, g and Supplementary Fig. 8), in contrast to identified dopaminergic neurons. None of the identified GABAergic neurons, and only 2 of 17 unidentified type II neurons, showed significant decreases in firing rate relative to when reward was delivered. None of the 11 type III neurons showed significant modulation by reward omission. Thus, the activity of type II and III neurons was modulated predominantly by reward-predicting cues but not actual reward.

Recent studies have revealed a diversity of responses of dopaminergic neurons to aversive stimuli: some are excited, others inhibited¹⁵. To test whether this diversity exists in dopaminergic and GABAergic VTA neurons, we delivered air puffs in some sessions. Identified dopaminergic neurons showed some diversity: although most significant responses were inhibitory, some were excitatory (Fig. 4h, i and Supplementary Fig. 9). In contrast, most type II and III neurons (and 13 of 14 identified GABAergic neurons) were excited by air puffs.

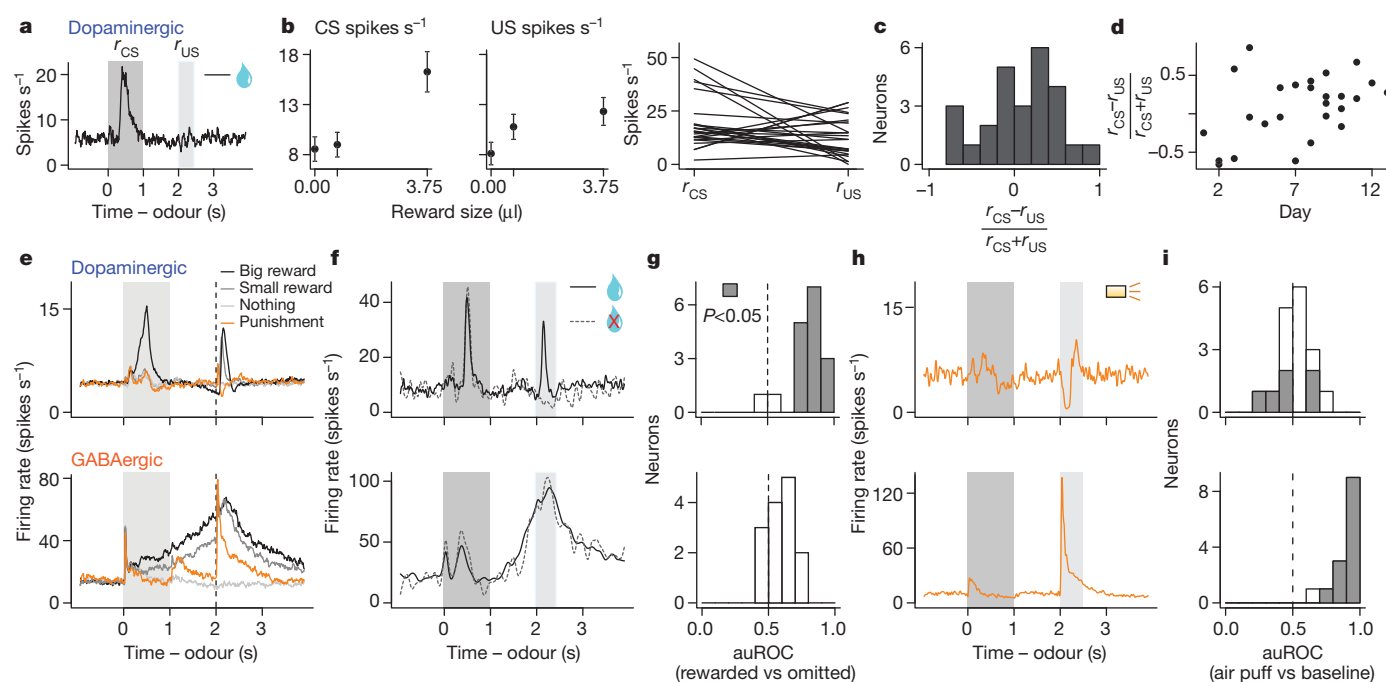


Figure 4 | Response variability based on CS-US preference, reward omission and air puffs. **a**, Response of a dopaminergic neuron during big-reward trials. **b**, Firing rate (mean \pm s.e.m.) versus reward size (left) and in response to big-reward-predicting CS and big-reward US for each dopaminergic neuron (right). **c**, Histogram of CS-US index for dopaminergic neurons. **d**, CS-US index versus day after the behaviour was learned. **e**, Average responses of dopaminergic and GABAergic neurons. **f**, Responses of a dopaminergic and GABAergic neuron for reward present (solid) and

unexpectedly absent (dashed) on big-reward trials. **g**, Histograms of differences in firing rates during the outcome period (2–2.5 s) between rewarded and reward-omitted trials for dopaminergic (top) and GABAergic (bottom) neurons. Values are represented using auROC (<0.5 , rewarded $<$ omitted; 0.5 , no difference; >0.5 , rewarded $>$ omitted). Significant values are filled (t -test, $P < 0.05$). **h**, Responses of a dopaminergic and GABAergic neuron during punishment trials. **i**, Histograms of auROC values during the air puff (2–2.5 s) relative to baseline (<0.5 , decrease; >0.5 , increase from baseline).

Detecting the discrepancy between expected and actual outcomes has a critical role in optimal learning^{1,22,23}. Although phasic firing of VTA dopaminergic neurons may act as such an error signal, how this is computed remains largely unknown. Models have postulated the existence of value-dependent, inhibitory input to dopaminergic neurons that persists during the delay between a CS and US (Supplementary Fig. 11a)^{1,23}. Our data indicate that VTA GABAergic neurons provide such an inhibitory input that counteracts excitatory drive from primary reward when the reward is expected. In addition, these neurons were excited by aversive stimuli, potentially contributing to suppression of firing in some dopaminergic neurons in response to aversive events (Fig. 4). Previous work showed that VTA GABAergic neurons receive inputs from prefrontal cortex and subcortical areas that could provide reward-related signals^{24–29}. Phasic excitation of VTA GABAergic neurons could be driven by inputs from lateral habenula neurons that are phasically excited by aversive stimuli²⁹. These habenular neurons do not show sustained activity between CS and US, so it is unlikely that they provide reward expectation signals to VTA GABAergic neurons. Instead, these signals may come from the pedunculo-pontine nucleus²⁵ or orbitofrontal cortex²⁷ (Supplementary Fig. 11b). VTA GABAergic neurons form synapses preferentially onto dendrites of dopaminergic neurons²⁸, whereas other inhibitory inputs form synapses onto their somata²⁹. Dendritic inhibition is thought to be weaker than somatic ‘shunting’ inhibition²⁸ but appears well suited for deriving graded outputs by ‘arithmetically’ combining excitatory and inhibitory inputs.

A major effect of drugs of addiction is inhibition of VTA GABAergic neurons^{7,8}. If VTA GABAergic neurons are involved in computation of RPE, inhibition of GABAergic neurons by addictive drugs could lead to sustained RPE even after the learned effects of drug intake are well established, thereby resulting in sustained reinforcement of drug taking³⁰. Understanding local circuits in VTA in the context of learning theory may thus provide crucial insights into normal as well as abnormal functions of reward circuits.

METHODS SUMMARY

All surgical and experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Harvard Institutional Animal Care and Use Committee. We injected male *DAT-Cre* and *Vgat-Cre* mice with adeno-associated virus carrying FLEX-ChR2 into the VTA and implanted a head plate and a microdrive containing six tetrodes and an optical fibre. While mice performed a classical conditioning task, we recorded spiking activity from VTA neurons. We delivered pulses of light to activate ChR2 and classified neurons as dopaminergic, GABAergic or unidentified. After experiments, we performed immunohistochemistry to localize recording sites amid dopaminergic neurons.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 May; accepted 2 December 2011.

Published online 18 January 2012.

- Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
- Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
- Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
- Swanson, L. W. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res. Bull.* **9**, 321–353 (1982).
- Margolis, E. B., Lock, H., Hjelmstad, G. O. & Fields, H. L. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol.* **577**, 907–924 (2006).
- Nair-Roberts, R. G. *et al.* Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience* **152**, 1024–1031 (2008).

- Hyman, S. E., Malenka, R. C. & Nestler, E. J. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci.* **29**, 565–598 (2006).
- Lüscher, C. & Malenka, R. C. Drug-evoked synaptic plasticity in addiction: from molecular changes to circuit remodeling. *Neuron* **69**, 650–663 (2011).
- Johnson, S. W. & North, R. A. Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J. Neurosci.* **12**, 483–488 (1992).
- Mansvelder, H. D., Keath, J. R. & McGehee, D. S. Synaptic mechanisms underlie nicotine-induced excitability of brain reward areas. *Neuron* **33**, 905–919 (2002).
- Szabo, B., Siemes, S. & Wallmichrath, I. Inhibition of GABAergic neurotransmission in the ventral tegmental area by cannabinoids. *Eur. J. Neurosci.* **15**, 2057–2061 (2002).
- Tan, K. R. *et al.* Neural bases for addictive properties of benzodiazepines. *Nature* **463**, 769–774 (2010).
- Dobi, A., Margolis, E. B., Wang, H.-L., Harvey, B. K. & Morales, M. Glutamatergic and nonglutamatergic neurons of the ventral tegmental area establish local synaptic contacts with dopaminergic and nondopaminergic neurons. *J. Neurosci.* **30**, 218–229 (2010).
- Steffensen, S. C., Svings, A. L., Pickel, V. M. & Henriksen, S. J. Electrophysiological characterization of GABAergic neurons in the ventral tegmental area. *J. Neurosci.* **18**, 8003–8015 (1998).
- Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
- Lammel, S. *et al.* Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**, 760–773 (2008).
- Nagel, G. *et al.* Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl Acad. Sci. USA* **100**, 13940–13945 (2003).
- Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8**, 1263–1268 (2005).
- Atasoy, D., Aponte, Y., Su, H. H. & Sternson, S. M. A FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J. Neurosci.* **28**, 7025–7030 (2008).
- Fiorillo, C. D., Newsome, W. T. & Schultz, W. The temporal precision of reward prediction in dopamine neurons. *Nature Neurosci.* **11**, 966–973 (2008).
- Takikawa, Y., Kawagoe, R. & Hikosaka, O. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J. Neurophysiol.* **92**, 2520–2529 (2004).
- Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. & Wagner, A. R.) 64–99 (New York, 1972).
- Houk, J. C., Adams, J. L. & Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 249–270 (MIT Press, 1995).
- Carr, D. B. & Sesack, S. R. Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *J. Neurosci.* **20**, 3864–3873 (2000).
- Okada, K., Toyama, K., Inoue, Y., Isa, T. & Kobayashi, Y. Different pedunculo-pontine tegmental neurons signal predicted and actual task rewards. *J. Neurosci.* **29**, 4858–4870 (2009).
- Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
- Takahashi, Y. K. *et al.* Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neurosci.* **14**, 1590–1597 (2011).
- Ormelchenko, N. & Sesack, S. R. Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Synapse* **63**, 895–906 (2009).
- Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B. & Holland, P. C. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61**, 786–800 (2009).
- Redish, A. D. Addiction as a computational process gone awry. *Science* **306**, 1944–1947 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Meister, V. N. Murthy, J. D. Schall and R. P. Heitz for comments, C. Dulac for sharing resources, C. I. Moore, J. Ritt and J. Siegle for advice about microdrives, K. Deisseroth for the AAV-FLEX-ChR2 construct, and E. Soucy and J. Greenwood for technical support. This work was supported by a Howard Hughes Medical Institute Fellowship from the Helen Hay Whitney Foundation (J.Y.C.); the Human Frontiers Science Program (S.H.); a Howard Hughes Medical Institute Collaborative Innovation Award, a Smith Family New Investigator Award, the Alfred Sloan Foundation, the Milton Fund (N.U.); F32 DK078478, P30 DK046200 (L.V.); and R01 DK075632, R01 DK089044, P30 DK046200, P30 DK057521 (B.B.L.).

Author Contributions J.Y.C. and S.H. collected and analysed data. J.Y.C., S.H. and N.U. designed experiments and wrote the paper. L.V. and B.B.L. generated *Vgat-Cre* mice.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.U. (uchida@mcb.harvard.edu).

METHODS

Animals. We used six adult male mice, backcrossed with C57/BL6J mice, heterozygous for Cre recombinase under the control of the *DAT* gene (B6.SJL-Slc6a3^{tm1.1(Cre)Bkmn}/J, The Jackson Laboratory)³¹ and six adult male mice, backcrossed with C57/BL6J mice, heterozygous for Cre recombinase under the control of the *Vgat* gene (*Vgat-ires-Cre*)³². Animals were housed on a 12-h dark/12-h light cycle (dark from 06:00 to 18:00) and each performed the conditioning task at the same time of day, between 07:00 and 18:00.

Surgery and viral injections. Mice were surgically implanted with a head plate and a microdrive containing electrodes and an optical fibre. During surgery, we injected 200–500 nl adeno-associated virus (AAV), serotype 5, carrying an inverted ChR2 (H134R)-EYFP flanked by double *loxP* sites^{18,19,33} into the VTA stereotactically (from bregma: 3.1 mm posterior, 0.8 mm lateral, 4–4.5 mm ventral). Our expression was highly selective (<1% of ChR2-positive neurons were negative for the dopaminergic neuron marker tyrosine hydroxylase, TH) and efficient (>90% of TH-expressing neurons co-expressed ChR2; Supplementary Figs 2 and 3).

For each unidentified type I neuron in *DAT*-Cre mice, we were unable to identify it as dopaminergic using our stringent criteria not because there was no response on the electrode after light pulses, but because the response did not match the shape of the spontaneous spike waveforms (open blue points in Fig. 3c). This may have been from changing the spike shape by opening many ChR2 or due to 'spillover' from nearby neurons, contaminating the signal.

After injections in *Vgat*-Cre mice, we confirmed that our ChR2 expression was selective (<1% of ChR2-positive neurons were negative for the GABAergic neuron marker glutamic acid decarboxylase, GAD65/67) and efficient (but less efficient than in *DAT*-Cre mice; >60% of GAD65/67-expressing neurons co-expressed ChR2) (Supplementary Fig. 2).

For fluorogold tracer injections (Fluorochrome), we injected two animals each with 10–25 nl into the following areas: prefrontal cortex (from bregma: 2.1 mm anterior, 0.25 mm lateral, 1.6 mm ventral), medial striatum (1.4 mm anterior, 1.0 mm lateral, 4.0 mm ventral) and basolateral amygdala (1.4 mm posterior, 2.8 mm lateral, 4.3 mm ventral).

All surgery was performed under aseptic conditions with animals under ketamine/medetomidine anaesthesia (60/0.5 mg kg⁻¹, intraperitoneal, respectively). Analgesia (ketoprofen, 5 mg kg⁻¹, intraperitoneal) was administered postoperatively.

Behavioural task. After >1 week of recovery, mice were water-deprived in their home cage. Weight was maintained above 90% of their full body weight. Animals were head-restrained using a custom-made metal plate and habituated for 1–2 d while head-restrained before training on the task. Odours were delivered with a custom-made olfactometer³⁴. Each odour was dissolved in paraffin oil at 1/100 dilution. Thirty microlitres of diluted odour was placed inside a filter-paper housing. Odours were isoamyl acetate, 1-butanol, *N*-butyl acetate, citral, eugenol, (+) limonene and (–) carvone, and differed for different animals. Odourized air was further diluted with filtered air by 1:10 to produce a 500 ml min⁻¹ total flow rate. Licks were detected by breaks of an infrared beam placed in front of the water tube.

We delivered one of four odours, selected pseudorandomly, for 1 s, followed by a delay of 1 s and an outcome. Each odour predicted a different outcome: a large drop of water (3.75 µl; valve open for 100 ms), a small drop of water (0.75 µl; valve open for 20 ms), no outcome, or an air puff delivered to the animal's face. Air puff trials were included for 17 identified dopaminergic neurons, 14 identified GABAergic neurons, and 15 type III neurons. Intertrial intervals (ITIs) were drawn from an exponential distribution with a rate parameter of 10 (that is, $ITI = 10\exp(-10x)$). This resulted in a flat ITI hazard function, ensuring that expectation about the start of the next trial did not increase over time. Data from *DAT*-Cre mice were obtained from 85 sessions (9–23 sessions per animal, 14 ± 4.8 sessions, mean \pm s.d.). For 17 identified dopaminergic neurons, 14 identified GABAergic neurons, and 11 type III neurons, we omitted rewards during 10% of big-reward trials. Animals performed between 400 and 1,000 trials per day (533 ± 120 trials, mean \pm s.d.). Data from *Vgat*-Cre mice were obtained from 71 sessions.

Electrophysiology. We recorded extracellularly from multiple neurons simultaneously using a custom-built 200-µm-fibreoptic-coupled screw-driven microdrive with six implanted tetrodes (four wires wound together). For three *Vgat*-Cre mice, we used a 105-µm fibre-optic. Tetrodes were glued to the fibre optic with epoxy. The ends of the tetrodes were 350–600 µm from the end of the fibre optic. Neural signals and time stamps for behaviour were recorded using a DigiLynx recording system (Neuralynx). Broadband signals from each wire filtered between 0.1 and 9,000 Hz were recorded continuously at 32 kHz. To extract the timing of spikes, signals were band-pass-filtered between 300–6,000 Hz and sorted online and offline using custom software.

To verify that our recordings targeted dopaminergic or GABAergic neurons, we used ChR2 to observe stimulation-locked spikes³⁵. The optical fibre was coupled with a diode-pumped solid-state laser with analogue amplitude modulation

(Laserglow Technologies). Before and after each behavioural session, we delivered trains of 10 light pulses, each 5-ms long, at 1, 5, 10, 20 and 50 Hz at 473 nm at 5–20 mW mm⁻². We used frequencies of 1, 10, 20, 50 and 100 Hz in *Vgat*-Cre mice. Higher intensities typically resulted in light-evoked spike waveforms that did not match spontaneous ones (open blue points in Fig. 3c and open red points in Fig. 3d). Therefore, we adjusted the light intensity after observing the responses. The increasing latency of light-evoked spiking as a function of stimulation frequency indicates that the response was not due to photochemical artefact (Supplementary Fig. 4f, j). Spike shape was measured using a broadband signal (0.1–9,000 Hz) sampled at 32 kHz. This ensured that particular features of the spike waveform were not missed. A previous study found decreasing probability of activation with increasing frequency³³. This discrepancy may be explained by the low levels of light we used to identify neurons and the difference in experimental preparation, *in vitro* versus *in vivo*³⁶.

We used two criteria to include a neuron in our data set. First, the neuron must have been well isolated (signal-to-noise ratio of >5 dB). Second, the neuron must have been recorded within 500 µm of an identified dopaminergic neuron (or a type I neuron in *Vgat*-Cre mice), to ensure that all neurons came from VTA. Using distance cutoffs of 400, 300 or 200 µm yielded similar estimates of the proportion of type I neurons as the full data set (49 of 91, 49 of 86 and 49 of 82, respectively, in *DAT*-Cre mice).

Recording sites were further verified histologically with electrolytic lesions using 15–20 s of 100 µA direct current and from the optical fibre track (Supplementary Fig. 12).

Data analysis. To measure firing rates, peristimulus time histograms (PSTHs) were constructed using 10-ms bins. To calculate spike density functions, PSTHs were convolved with a function resembling a postsynaptic potential³⁷. To determine whether a neuron showed a significant task-related response, we calculated an ANOVA on the trial-by-trial firing rates (spikes s⁻¹) during the baseline period (1 s before odour onset), CS period (from odour onset to odour offset), delay (from odour offset to outcome onset), and US period (from outcome onset to 500 ms after outcome onset). The factors were task epoch (that is, baseline, CS, delay, or US) and outcome type. All neurons showed a main effect of time; 183 of 187 neurons showed a main effect of outcome type; 184 of 187 neurons showed a time–outcome interaction ($P < 0.001$).

Light-evoked spikes were detected during the 10 ms after light onset. If <20% of light pulses had a spike (defined as a waveform that matched that of the isolated unit) during the 10 ms after light onset (lower left points in Fig. 3c, d), the maximum voltage during that interval was used as the light-evoked 'response'. Cross-correlations between spontaneous and light-evoked spike waveforms were calculated by aligning the positive peak of each waveform, averaging separately, and aligning the peaks of the averages. The cross-correlation coefficients reported in Fig. 3 are from a lag of 0 ms. The correlation was calculated using the full duration of the spontaneous spike (spike duration was measured as the first time until the last time at which the voltage was significantly different from baseline using Wilcoxon rank sum tests). The energy of the light-evoked waveform is defined as the integral of the squared voltage values ($\int v^2 dt$). To ensure that measurements of the fidelity of light-evoked responses were not biased by poorly isolated neurons, we plotted the probability of firing at 50 Hz light stimulation against L-ratio, a measure of isolation quality³⁸ (Supplementary Fig. 13). L-ratio approximates the fraction of 'contaminated' spikes; smaller L-ratios indicate better isolation. One hundred and eighty of 187 neurons in the data set had L-ratios <0.05.

CS–US indices were calculated as $(CS - US)/(CS + US)$, where CS is the difference between the peak firing rate (maximum value of the PSTH) in the 1 s after odour onset and the baseline firing rate, and US is the difference between the peak firing rate in the 0.5 s after reward onset and the baseline firing rate. The baseline firing rate was calculated as the mean of the PSTH in the 0.5 s before odour onset.

ROC curves were calculated by comparing the distribution of firing rates across trials in 100-ms bins to the distribution of baseline firing rates (900 ms before odour onset; Supplementary Fig. 1). PCA was calculated using the singular value decomposition of the auROC. Learning was measured using the lick responses (mean licks s⁻¹) across consecutive days of behavioural sessions. Curves were fit with logistic functions, $k/(1 + \exp(-t))$. The task was considered fully learned if the learning rate was within 95% of the carrying capacity of the logistic (k). All 12 animals learned the task within the first 3 days.

Hierarchical clustering was done using the first three PCs of the auROC curves using a Euclidean distance metric and complete agglomeration method.

All statistical tests were done with corrections for multiple comparisons (Bonferroni or Tukey). Analyses were done with R (<http://www.r-project.org/>).

Immunohistochemistry. After recording, which lasted between 9 and 23 days, mice were given an overdose of ketamine/medetomidine, exsanguinated with saline, perfused with paraformaldehyde, and brains were cut in 50 µm coronal or horizontal sections. Sections were immunostained with antibodies to TH and

secondary antibodies labelled with Cy3 (Jackson Immunoresearch). Sections were further stained with 4',6-diamidino-2-phenylindole (DAPI) or TO-PRO-3 (Invitrogen) to visualize nuclei. Recording sites were identified and verified to be amid EYFP staining and TH staining in VTA. Brain slices from additional virus-injected *DAT*-Cre and *Vgat*-Cre animals were stained for TH and GAD-65/67 (Millipore) with secondary antibodies labelled with Cy3 or Alexa 594 (Invitrogen). Cellular fluorescence intensities of ChR2-EYFP were measured using Volocity Image Analysis Software (Perkin Elmer).

31. Bäckman, C. M. *et al.* Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. *Genesis* **44**, 383–390 (2006).
32. Vong, L. *et al.* Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).
33. Tsai, H. C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080–1084 (2009).
34. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nature Neurosci.* **6**, 1224–1229 (2003).
35. Lima, S. Q., Hromádka, T., Znamenskiy, P. & Zador, A. M. PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS ONE* **4**, e6099 (2009).
36. Zhao, S. *et al.* Cell-type specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nature Methods* **8**, 745–752 (2011).
37. Thompson, K. G., Hanes, D. P., Bichot, N. P. & Schall, J. D. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J. Neurophysiol.* **76**, 4040–4055 (1996).
38. Schmitzer-Torbert, N. & Redish, A. D. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the Multiple T Task. *J. Neurophysiol.* **91**, 2259–2272 (2004).

Vaccine protection against acquisition of neutralization-resistant SIV challenges in rhesus monkeys

Dan H. Barouch^{1,2}, Jinyan Liu¹, Hualin Li¹, Lori F. Maxfield¹, Peter Abbink¹, Diana M. Lynch¹, M. Justin Iampietro¹, Adam SanMiguel¹, Michael S. Seaman¹, Guido Ferrari³, Donald N. Forthal⁴, Ilnour Ourmanov⁵, Vanessa M. Hirsch⁵, Angela Carville⁶, Keith G. Mansfield⁶, Donald Stablein⁷, Maria G. Pau⁸, Hanneke Schuitemaker⁸, Jerald C. Sadoff⁸, Erik A. Billings⁹, Mangala Rao⁹, Merlin L. Robb⁹, Jerome H. Kim⁹, Mary A. Marovich⁹, Jaap Goudsmit^{8*} & Nelson L. Michael^{9*}

Preclinical studies of human immunodeficiency virus type 1 (HIV-1) vaccine candidates have typically shown post-infection virological control, but protection against acquisition of infection has previously only been reported against neutralization-sensitive virus challenges^{1–3}. Here we demonstrate vaccine protection against acquisition of fully heterologous, neutralization-resistant simian immunodeficiency virus (SIV) challenges in rhesus monkeys. Adenovirus/poxvirus and adenovirus/adenovirus-vector-based vaccines expressing SIV_{SME543} Gag, Pol and Env antigens resulted in an 80% or greater reduction in the per-exposure probability of infection^{4,5} against repetitive, intrarectal SIV_{MAC251} challenges in rhesus monkeys. Protection against acquisition of infection showed distinct immunological correlates compared with post-infection virological control and required the inclusion of Env in the vaccine regimen. These data demonstrate the proof-of-concept that optimized HIV-1 vaccine candidates can block acquisition of stringent, heterologous, neutralization-resistant virus challenges in rhesus monkeys.

Despite the recent demonstration of partial HIV-1 vaccine efficacy in humans⁶, the immune responses required to protect against acquisition of infection have remained unclear. Preclinical studies of HIV-1 vaccine candidates have begun to elucidate immunological correlates of protection against neutralization-sensitive viruses^{1–3}, but no study has to date reported vaccine protection against acquisition of heterologous, neutralization-resistant virus challenges^{1,7,8}. Mucosal SIV_{MAC251} infection of rhesus monkeys represents a stringent preclinical model of a highly pathogenic, neutralization-resistant virus swarm^{1,9,10}, and repetitive mucosal challenges more closely mimic sexual HIV-1 transmission in humans than do single high-dose challenges¹⁰. We therefore performed two studies to evaluate the protective efficacy of optimized adenovirus/poxvirus and adenovirus/adenovirus vector-based vaccines against repetitive, heterologous, intrarectal SIV_{MAC251} challenges in rhesus monkeys.

In the first study, 40 Indian-origin rhesus monkeys (*Macaca mulatta*) that did not express the class I alleles *Mamu-A*01*, *Mamu-B*08*, and *Mamu-B*17* associated with spontaneous virological control^{11–13} were immunized by the intramuscular route with the following vaccine regimens expressing SIV_{SME543} Gag-Pol and Env immunogens ($N = 8$ per group): (1) DNA prime, modified vaccinia Ankara (MVA) boost; (2) MVA prime, MVA boost; (3) adenovirus serotype 26 (Ad26) prime, MVA boost; (4) MVA prime, Ad26 boost; and (5) sham control. Groups were balanced for susceptible and resistant *TRIM5 α* alleles^{1,14}. Monkeys were primed once at week 0 with 2×10^{10} viral particles of Ad26 vectors or 10^8 plaque-forming units (p.f.u.) of MVA vectors, or three times at weeks 0, 4, and 8 with 5 mg of DNA vaccines. Animals were then boosted once at week 24 with 2×10^{10} viral particles of Ad26 vectors or 10^8 p.f.u. of MVA vectors.

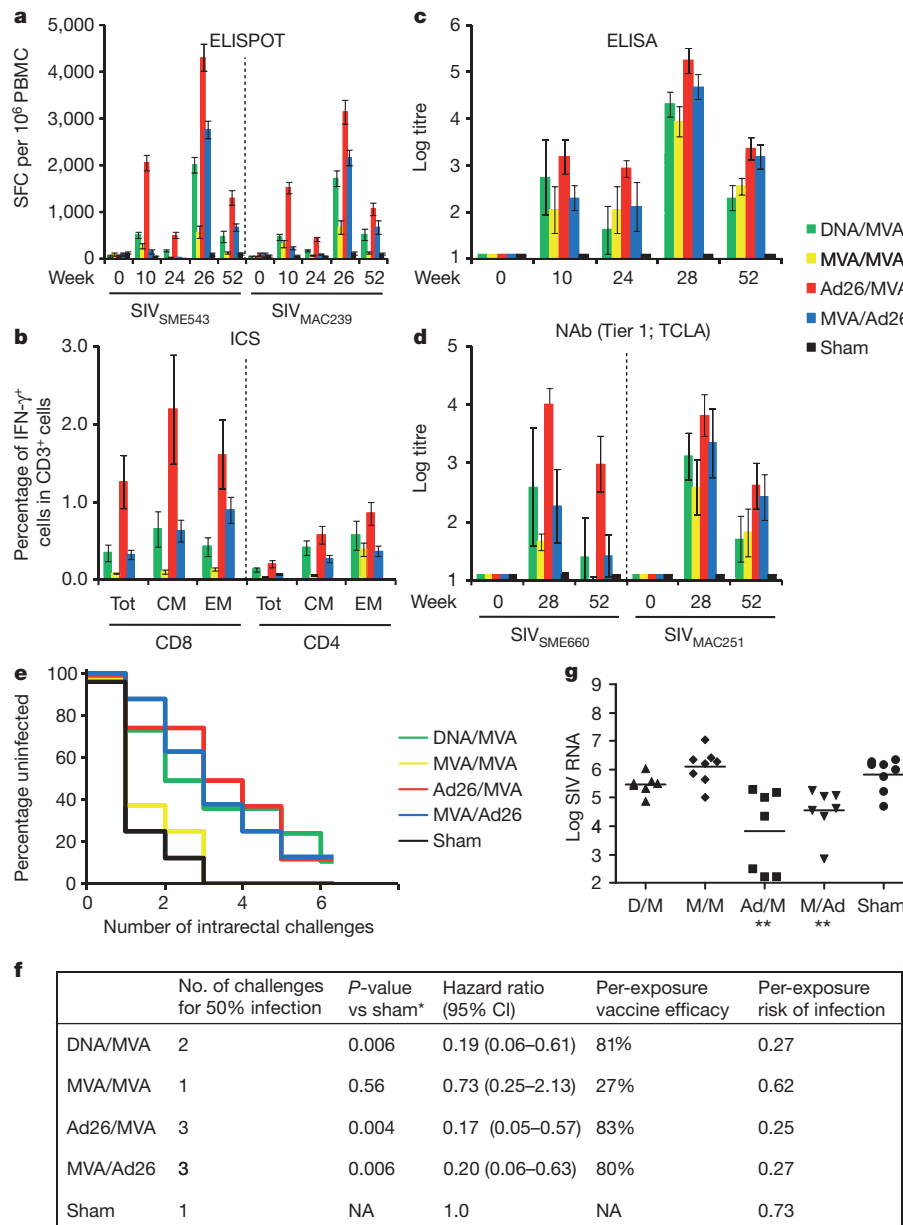
The vaccine regimens elicited distinct profiles of cellular and humoral immune responses, as measured by IFN- γ ELISPOT assays (Fig. 1a and Supplementary Fig. 1), multiparameter intracellular cytokine staining (ICS) assays^{8,15–17} (Fig. 1b and Supplementary Fig. 2), cellular immune breadth (Supplementary Fig. 3), SIV_{MAC251} Env-specific binding antibody ELISAs (Fig. 1c), tier 1 neutralizing antibody (NAb) assays against tissue culture laboratory adapted (TCLA) tier 1 SIV_{SME660} (CP3C-P-A8) and SIV_{MAC251} (TCLA) pseudoviruses (Fig. 1d), and antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cell-mediated virus inhibition (ADCVI) assays (Supplementary Fig. 4). Tier 2 NAb responses against neutralization-resistant SIV_{SME660} (CR54-PK-2A5) and SIV_{MAC251} (SIV_{MAC251}-30) pseudoviruses, however, were below the 50% neutralization cutoff for positivity, although positive trends were observed in all vaccinated groups (Supplementary Fig. 4).

To evaluate the protective efficacy of these vaccine regimens, all monkeys were challenged repetitively beginning at week 52 (six months following the boost immunization) with six intrarectal inoculations of the heterologous virus SIV_{MAC251} using a 1:1,000 dilution (930 half-maximal tissue-culture infectious dose (TCID₅₀)) of our challenge stock⁹. After the first challenge, 75% of sham control monkeys became infected, compared with only 12–25% of the animals that received the heterologous vector regimens DNA/MVA, Ad26/MVA, and MVA/Ad26 (Fig. 1e). The percentage of uninfected animals declined proportionately with each challenge, and the majority of vaccinees and all controls were infected by the end of the challenge protocol. Monkeys that received the Ad26/MVA and MVA/Ad26 vaccines required three challenges to infect 50% of animals in each group, whereas only one challenge was required to infect 50% of animals in the control group ($P = 0.004$ and $P = 0.006$, respectively, Wald tests, proportional hazard model). The heterologous vector regimens also showed decreased hazard ratios of 0.17 (95% confidence interval (CI) 0.05–0.57) to 0.20 (CI 0.06–0.63) compared with the controls, corresponding to an 80–83% reduction in the per-exposure probability of infection (Fig. 1f; vaccine efficacy $VE = 1 - \text{hazard ratio}$), as previously described^{4,5}. These data demonstrate vaccine protection against acquisition of infection following repetitive, heterologous, intrarectal SIV_{MAC251} challenges.

Control monkeys showed peak viral loads on day 14 following infection and then relatively stable mean set point viral loads of 5.85 log SIV RNA copies per millilitre (Supplementary Fig. 5). The Ad26/MVA and the MVA/Ad26 vaccines resulted, respectively, in at least 2.32 and 1.08 log reductions of mean set point viral loads compared with sham controls for over 250 days ($P = 0.0037$ for each vaccine versus sham, Wilcoxon rank-sum tests) (Fig. 1g and Supplementary Fig. 5). Moreover, half the animals in the Ad26/MVA

¹Division of Vaccine Research, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA. ²Ragon Institute of MGH, MIT, and Harvard, Boston, Massachusetts 02114, USA. ³Duke University Medical Center, Durham, North Carolina 27710, USA. ⁴University of California Irvine School of Medicine, Irvine, California 92697, USA. ⁵National Institute of Allergy and Infectious Diseases, Bethesda, Massachusetts 20892, USA. ⁶New England Primate Research Center, Southborough, Massachusetts 01772, USA. ⁷EMMES Corporation, Rockville, Maryland 20850, USA. ⁸Crucell Holland BV, 2301 CA, Leiden, The Netherlands. ⁹US Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, Maryland 20910, USA.

*These authors contributed equally to this work.



*Wald test, proportional hazard model

Figure 1 | Immunogenicity and protective efficacy of the adenovirus/poxvirus vaccines. **a**, Cellular immune responses to SIV_{SME543} and SIV_{MAC239} Gag, Pol and Env as determined by IFN- γ ELISPOT assays at weeks 0, 10, 24, 26, and 52. PBMC, peripheral blood mononuclear cells; SFC, spot-forming cells. **b**, CD8⁺ and CD4⁺ total, central/transitional memory (CM; CD28⁺CD95⁺), and effector memory (EM; CD28⁺CD95⁺) responses to Gag, Pol and Env as determined by multiparameter IFN- γ ICS assays at week 26. **c**, SIV_{MAC251} Env ELISAs at weeks 0, 10, 24, 28, and 52. **d**, SIV_{SME660} and SIV_{MAC251} tier 1 pseudovirus NAb assays at weeks 0, 28, and 52. Error bars

group either demonstrated rapid and durable virological control to undetectable levels (Fig. 1g; $N = 3$) or remained uninfected (Fig. 1e; $N = 1$). The Ad26/MVA and MVA/Ad26 vaccines also afforded a survival advantage as compared with the controls ($P = 0.025$, log-rank test) (Supplementary Fig. 6).

We next evaluated the immunological correlates of protection against acquisition of infection, defined as the number of challenges required to establish infection, and virological control, defined as set point viral loads. Our pre-specified primary immunological correlates analysis (Supplementary Table 1) demonstrated that protection against acquisition of infection was best correlated with Env-specific binding ELISA antibody responses (Fig. 2a; $P < 0.0001$, Spearman

represent s.e.m. **e**, Number of challenges required for acquisition of infection in each vaccine group. **f**, Statistical analyses include the number of challenges required for 50% infection, hazard ratios with 95% confidence intervals (CI), per-exposure vaccine efficacy and per-exposure risks of infection in each group. P -values reflect Wald tests using a proportional hazard model. **g**, Log SIV RNA copies per ml are depicted for each vaccine group at viral set point (day 84). ** $P = 0.0037$, Wilcoxon rank-sum tests. The horizontal lines represent mean set point log viral loads. Ad, adenovirus; D, DNA; M, MVA.

rank-correlation test) and tier 1 NAb titres (Fig. 2b; $P = 0.0034$) immediately before challenge. Protection against acquisition of infection also correlated with V2-specific antibodies that presumably represented a subset of total Env-specific binding antibodies (Fig. 2e, f; $P < 0.0001$). Virological control was correlated with Gag ELISPOT breadth (Fig. 2c; $P = 0.0002$) and magnitude (Fig. 2d; $P = 0.0058$) immediately before challenge, consistent with our previous observations¹⁸.

In our exploratory immunological correlates analysis, we evaluated 35 humoral and cellular immune parameters at both peak and memory time points before challenge as possible immunological correlates of acquisition and virological control following challenge. No additional immune parameters were significantly correlated with protection

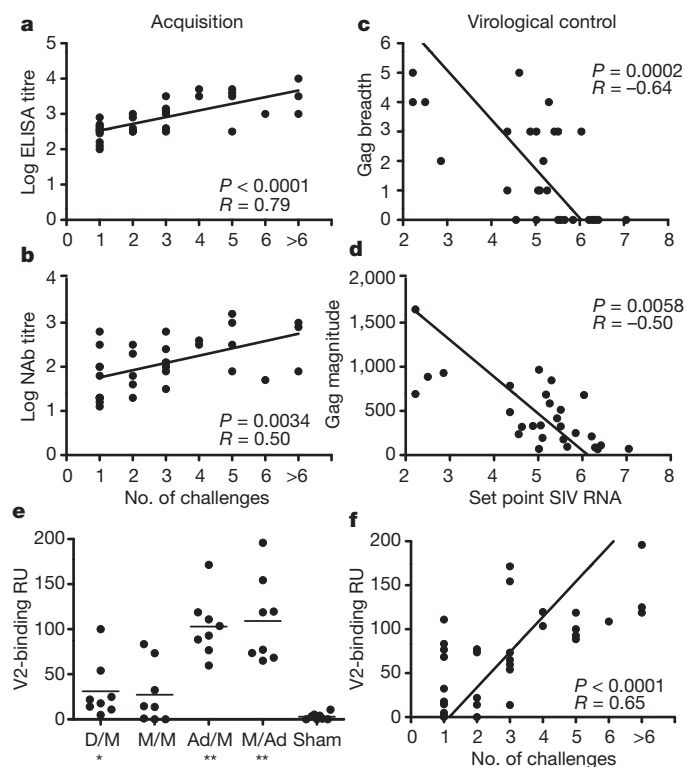


Figure 2 | Correlates of protection against acquisition of infection and virological control with the adenovirus/poxvirus vaccines. **a, b,** Correlation of log ELISA titres immediately before challenge (**a**) and log tier 1 NAb titres immediately before challenge (**b**) with the number of challenges required to establish infection. **c, d,** Correlation of Gag ELISPOT breadth before challenge (**c**) and Gag ELISPOT magnitude before challenge (**d**) with set point viral loads following challenge. Correlates analyses included the 32 vaccinated monkeys (**a, b**) or the 29 vaccinated animals that became infected (**c, d**) and did not include the sham controls. P -values reflect Spearman rank-correlation tests. **e,** V2-specific binding antibodies assessed by surface plasmon resonance response units (RU) for each vaccine group at week 30. $*P = 0.002$, $**P = 0.0007$, Wilcoxon rank-sum tests. The horizontal lines represent mean responses. **f,** Correlation of V2-specific antibody responses with the number of challenges required to establish infection.

against acquisition of infection in this analysis after multiple comparison adjustments (Supplementary Table 2). Gag-, Pol- and Env-specific effector memory CD8⁺ T-lymphocyte responses exhibited trends towards protection against acquisition, but did not achieve statistical significance according to our pre-specified criteria. In contrast, multiple humoral and cellular immune responses were significantly correlated with virological control (Supplementary Table 3), including Env ELISA, NAb and ADCC responses as well as Gag ELISPOT magnitude and breadth, Pol ELISPOT magnitude and Env CD4⁺ effector memory responses. These data support a model in which protection against acquisition of infection is correlated with vaccine-elicited Env antibody responses, whereas virological control is correlated with both T-lymphocyte and antibody responses. These distinct immunological correlates probably reflect fundamentally different biologic requirements for blocking establishment of infection at the mucosal site of inoculation compared with controlling viral replication after infection has become disseminated¹⁹. However, the actual mechanisms of protection remain to be determined.

We next evaluated directly the hypothesis that Env was critical for blocking acquisition of infection in this system. In the second study, 40 rhesus monkeys that did not express the class I alleles *Mamu-A*01*, *Mamu-B*08* and *Mamu-B*17* were immunized by the intramuscular route with Ad35 prime²⁰, Ad26 boost²¹ vaccine regimens expressing (1) SIV_{SME543} Gag-Pol ($N = 16$), (2) SIV_{SME543} Gag-Pol and Env ($N = 16$) and (3) sham controls ($N = 8$). Groups were balanced for

susceptible and resistant *TRIM5 α* alleles^{1,14}. Monkeys were primed once at week 0 with 2×10^{10} viral particles of Ad35 vectors and boosted once at week 24 with 2×10^{10} viral particles of Ad26 vectors. Cellular immune responses were assessed by IFN- γ ELISPOT assays (Fig. 3a and Supplementary Fig. 7) and multiparameter ICS assays in both the periphery (Fig. 3b and Supplementary Fig. 8) and in colorectal mucosa (Supplementary Fig. 9). Env-specific humoral immune responses were assessed by ELISAs in both the periphery (Fig. 3c) and in colorectal mucosa (Supplementary Fig. 10), tier 1 NAb assays (Fig. 3d) and ADCC assays (Supplementary Fig. 11). Only marginal tier 2 NAb responses were observed (Supplementary Fig. 11).

We assessed protective efficacy of these vaccine regimens against repetitive, heterologous, intrarectal SIV_{MAC251} challenges as described in the first study. After the first challenge, 50% of sham control monkeys became infected, compared with only 12% of the animals that received the Gag-Pol-Env vaccine (Fig. 3e). The monkeys that received the Gag-Pol-Env vaccine required four challenges to infect 50% of animals in each group, whereas only one challenge was required to infect 50% of animals in the control group (Fig. 3f; $P = 0.002$, Wald test, proportional hazard model). Moreover, the Gag-Pol-Env vaccine resulted in a decreased hazard ratio of 0.20 (CI 0.07–0.55), corresponding to an 80% reduction in the per-exposure probability of infection. In contrast, the Gag-Pol vaccine afforded only a marginal protective effect, demonstrating the critical role of Env in blocking acquisition of infection in this model. The Gag-Pol and Gag-Pol-Env vaccines resulted in, respectively, at least 1.59 log and 2.18 log reductions of set point viral loads compared with controls (Fig. 3g and Supplementary Fig. 12; $P = 0.0006$ and 0.0002 , respectively, Wilcoxon rank-sum tests). Immunological correlates of protection against acquisition of infection were consistent with the first study, and both peripheral (Fig. 4a–c) and rectal (Fig. 4d) Env-specific IgG correlated with reduced acquisition risk.

Taken together, these data demonstrate that optimized adenovirus/poxvirus and adenovirus/adenovirus vector-based vaccines afforded significant protection against acquisition of infection following highly pathogenic, heterologous, neutralization-resistant SIV_{MAC251} challenges in rhesus monkeys (Figs 1e, 3e and Supplementary Fig. 13). Although several studies have previously shown partial protection against acquisition of neutralization-sensitive virus challenges^{1–3}, no HIV-1 vaccine candidate has to date blocked acquisition of heterologous, difficult-to-neutralize virus challenges, including Ad5 (ref. 7), DNA/Ad5 (ref. 1) and cytomegalovirus⁸ vaccines. In particular, a recent study demonstrated that a DNA/Ad5 vaccine afforded partial protection against acquisition of SIV_{SME660}, which is a neutralization-sensitive tier 1A virus in TZM-bl neutralization assays, but the same vaccine afforded no efficacy against neutralization-resistant SIV_{MAC251} (ref. 1), highlighting important differences in the stringencies between these two SIV challenge models as well as potentially important phenotypic differences between adenovirus serotypes¹⁷. However, we note that the acquisition effect in the present study was relative rather than absolute, and that the majority of vaccinees became infected by the end of the challenge protocol.

Our studies also demonstrate that inclusion of Env in the vaccine was required for the acquisition effect (Fig. 3e), despite an 18% difference in the Env amino acid sequences between the vaccine strain and challenge virus. Moreover, our immunological correlates analyses (Figs 2, 4 and Supplementary Tables 1–3) suggest that Env-specific antibodies are critical for blocking acquisition of infection, whereas multiple cellular and humoral immune responses correlate with virological control, although the actual mechanisms of protection remain to be determined. In addition, the RV144 immunological correlates analyses raised the hypothesis that vaccine-elicited V1/V2-specific antibodies may reduce HIV-1 acquisition risk in humans²². Our data (Figs 2f, 4c) are consistent with this hypothesis, although it remains unclear whether V2-specific antibodies actually protect or simply represent a marker for other Env-specific antibodies or other protective factors.

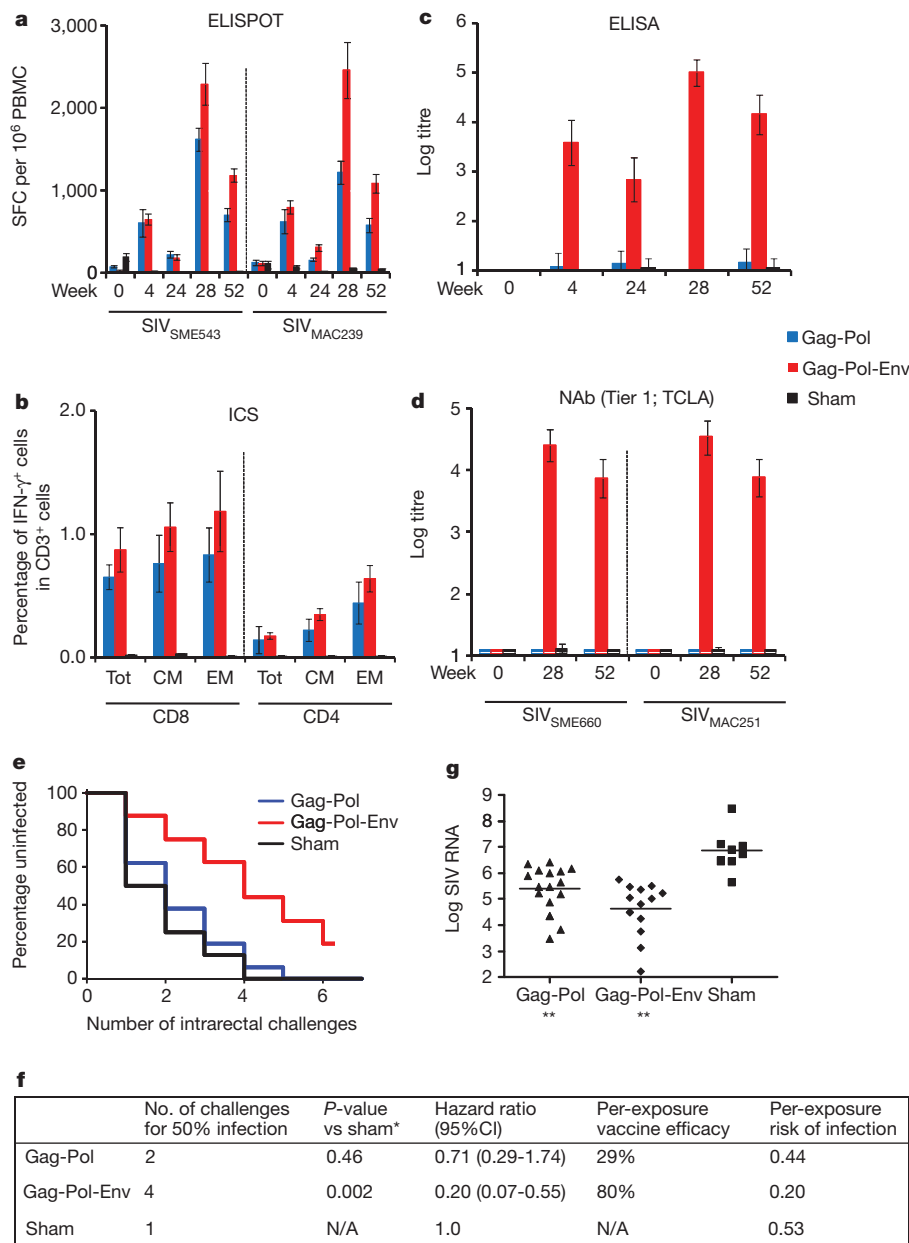


Figure 3 | Immunogenicity and protective efficacy of the adenovirus/adenovirus vaccines. **a**, Cellular immune responses to SIV_{SME543} and SIV_{MAC239} Gag, Pol and Env as determined by IFN- γ ELISPOT assays at weeks 0, 4, 24, 28 and 52. **b**, CD8⁺ and CD4⁺ total, central/transitional memory (CM; CD28⁺CD95⁺), and effector memory (EM; CD28⁺CD95⁺) responses to Gag, Pol and Env as determined by multiparameter IFN- γ ICS assays at week 28. **c**, SIV_{MAC251} Env ELISAs at weeks 0, 4, 24, 28 and 52. **d**, SIV_{SME660} and SIV_{MAC251} tier 1 pseudovirus NAb assays at weeks 0, 28 and 52. Error bars

represent s.e.m. **e**, Number of challenges required for acquisition of infection in each vaccine group. **f**, Statistical analyses include the number of challenges required for 50% infection, hazard ratios with 95% confidence intervals (CI), per-exposure vaccine efficacy, and per-exposure risks of infection in each group. *P*-values reflect Wald tests using a proportional hazard model. **g**, Log SIV RNA copies per ml are depicted for each vaccine group at viral set point (day 84). ***P* < 0.001, Wilcoxon rank-sum tests. The horizontal lines reflect mean set point log viral loads.

Considerable efforts are currently underway to identify and to reverse engineer potent, broadly reactive monoclonal antibodies^{23,24}. Although the induction of such NAb responses by a vaccine would presumably be highly desirable, no Env immunogens have to date been developed that can elicit these responses. Our findings suggest that a substantial degree of protection can be achieved against stringent virus challenges even in the absence of high titres of tier 2 NABs, perhaps reflecting the importance of antibody effector functions that may not be fully measured by traditional virus neutralization assays. Of note, the partial protection in the present study was observed with vectored

Env and without a purified Env protein subunit boost. The degree to which an Env protein boost may further improve the protective efficacy afforded by these vaccine regimens remains to be determined.

In summary, our data demonstrate the proof-of-concept that vaccination can protect against acquisition of stringent, heterologous, neutralization-resistant SIV_{MAC251} challenges in rhesus monkeys. These findings, together with the observations of a critical requirement for Env and the distinct immunological correlates of protection against acquisition of infection and virological control, pave novel paths forward for HIV-1 vaccine development.

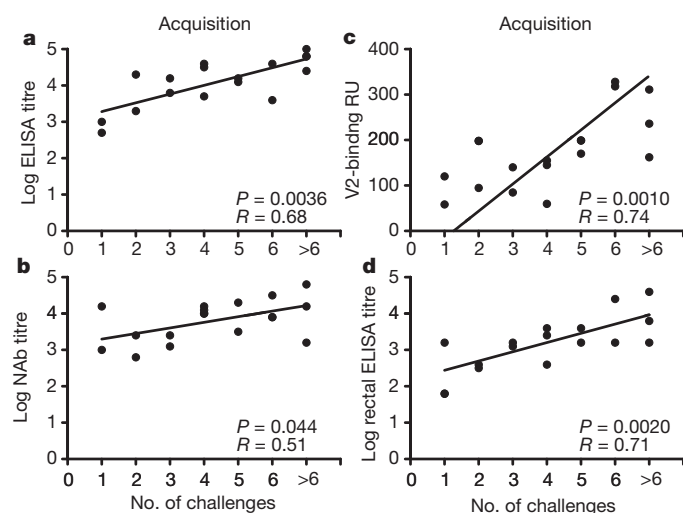


Figure 4 | Correlates of protection against acquisition of infection with the adenovirus/adenovirus vaccines. a–d, Correlation of log ELISA titres immediately before challenge (a), log tier 1 NAb titres immediately before challenge (b), V2-specific antibody responses (c) and rectal IgG antibody responses (d) with the number of challenges required to establish infection. Correlates analyses included the 16 Gag-Pol-Env-vaccinated monkeys and did not include the Gag-Pol-vaccinated monkeys or the sham controls. *P*-values reflect Spearman rank-correlation tests.

METHODS SUMMARY

For each study, 40 Indian-origin rhesus monkeys (*Macaca mulatta*) were vaccinated with DNA, MVA²⁵, Ad26 (ref. 21) or Ad35 (ref. 20) expressing SIV_{SME543} Gag-Pol and/or Env antigens or received a sham vaccine. Cellular immune responses were measured by IFN- γ ELISPOT¹⁸, multiparameter intracellular cytokine staining (ICS)^{8,15–17} and epitope mapping²⁶ assays. Humoral immune responses were measured by Env ELISA²⁷, TZM-bl pseudovirus neutralizing antibody (NAb)²⁸, antibody-dependent cellular cytotoxicity (ADCC)²⁹ and antibody-dependent cell-mediated virus inhibition (ADCVI)³⁰ assays. Six months after the boost immunization, all monkeys received six challenges by the intrarectal route with the fully heterologous, neutralization-resistant virus SIV_{MAC251} using a 1:1,000 dilution (930 TCID₅₀) of our challenge stock⁹. Protective efficacy was determined by resistance to acquisition of infection, defined as the number of challenges required to establish infection, and virological control, defined as set point viral loads. Immunological correlates of protection against acquisition of infection and virological control were evaluated by pre-specified primary and exploratory analyses.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 September; accepted 7 December 2011.

Published online 4 January; corrected 1 February 2012 (see full-text HTML version for details).

- Letvin, N. L. *et al.* Immune and genetic correlates of vaccine protection against mucosal infection by SIV in monkeys. *Sci. Transl. Med.* **3**, 81ra36 (2011).
- Lai, L. *et al.* Prevention of infection by a granulocyte-macrophage colony-stimulating factor co-expressing DNA/modified vaccinia Ankara simian immunodeficiency virus vaccine. *J. Infect. Dis.* **204**, 164–173 (2011).
- Barnett, S. W. *et al.* Antibody-mediated protection against mucosal simian immunodeficiency virus challenge of macaques immunized with alphavirus replicon particles and boosted with trimeric envelope glycoprotein in MF59 adjuvant. *J. Virol.* **84**, 5975–5985 (2010).
- Hudgens, M. G. *et al.* Power to detect the effects of HIV vaccination in repeated low-dose challenge experiments. *J. Infect. Dis.* **200**, 609–613 (2009).
- Hudgens, M. G. & Gilbert, P. B. Assessing vaccine effects in repeated low-dose challenge experiments. *Biometrics* **65**, 1223–1232 (2009).
- Rerks-Ngarm, S. *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N. Engl. J. Med.* **361**, 2209–2220 (2009).
- Wilson, N. A. *et al.* Vaccine-induced cellular immune responses reduce plasma viral concentrations after repeated low-dose challenge with pathogenic simian immunodeficiency virus SIVmac239. *J. Virol.* **80**, 5875–5885 (2006).
- Hansen, S. G. *et al.* Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* **473**, 523–527 (2011).

- Liu, J. *et al.* Low-dose mucosal simian immunodeficiency virus infection restricts early replication kinetics and transmitted virus variants in rhesus monkeys. *J. Virol.* **84**, 10406–10412 (2010).
- Keele, B. F. *et al.* Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. *J. Exp. Med.* **206**, 1117–1134 (2009).
- Yant, L. J. *et al.* The high-frequency major histocompatibility complex class I allele Mamu-B*17 is associated with control of simian immunodeficiency virus SIVmac239 replication. *J. Virol.* **80**, 5074–5077 (2006).
- Mothé, B. R. *et al.* Expression of the major histocompatibility complex class I molecule Mamu-A*01 is associated with control of simian immunodeficiency virus SIVmac239 replication. *J. Virol.* **77**, 2736–2740 (2003).
- Loffredo, J. T. *et al.* Mamu-B*08-positive macaques control simian immunodeficiency virus replication. *J. Virol.* **81**, 8827–8832 (2007).
- Lim, S. Y. *et al.* TRIM5 α modulates immunodeficiency virus control in rhesus monkeys. *PLoS Pathog.* **6**, e1000738 (2010).
- Okoye, A. *et al.* Progressive CD4⁺ central memory T cell decline results in CD4⁺ effector memory insufficiency and overt disease in chronic SIV infection. *J. Exp. Med.* **204**, 2171–2185 (2007).
- Pitcher, C. J. *et al.* Development and homeostasis of T cell memory in rhesus macaque. *J. Immunol.* **168**, 29–43 (2002).
- Liu, J. *et al.* Magnitude and phenotype of cellular immune responses elicited by recombinant adenovirus vectors and heterologous prime-boost regimens in rhesus monkeys. *J. Virol.* **82**, 4844–4852 (2008).
- Liu, J. *et al.* Immune control of an SIV challenge by a T-cell-based vaccine in rhesus monkeys. *Nature* **457**, 87–91 (2009).
- Haase, A. T. Targeting early infection to prevent HIV-1 mucosal transmission. *Nature* **464**, 217–223 (2010).
- Vogels, R. *et al.* Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. *J. Virol.* **77**, 8263–8271 (2003).
- Abbink, P. *et al.* Comparative seroprevalence and immunogenicity of six rare serotype recombinant adenovirus vaccine vectors from subgroups B and D. *J. Virol.* **81**, 4654–4663 (2007).
- Haynes, B. Case control study of the RV144 trial for immune correlates: the analysis and way forward. *AIDS Vaccine Conference 2011* (Bangkok, Thailand, 2011).
- Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285–289 (2009).
- Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856–861 (2010).
- Ourmanov, I. *et al.* Comparative efficacy of recombinant modified vaccinia virus Ankara expressing simian immunodeficiency virus (SIV) Gag-Pol and/or Env in macaques challenged with pathogenic SIV. *J. Virol.* **74**, 2740–2751 (2000).
- Barouch, D. H. *et al.* Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nature Med.* **16**, 319–323 (2010).
- Nkolola, J. P. *et al.* Breadth of neutralizing antibodies elicited by stable, homogeneous clade A and clade C HIV-1 gp140 envelope trimers in guinea pigs. *J. Virol.* **84**, 3270–3279 (2010).
- Montefiori, D. C. Evaluating neutralizing antibodies against HIV, SIV and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* **12**, Unit 12.11 (2004).
- Ferrari, G. *et al.* An HIV-1 gp120 envelope human monoclonal antibody that recognizes a C1 conformational epitope mediates potent antibody-dependent cellular cytotoxicity (ADCC) activity and defines a common ADCC epitope in human HIV-1 serum. *J. Virol.* **85**, 7029–7036 (2011).
- Forthal, D. N., Landucci, G. & Daar, E. S. Antibody from patients with acute human immunodeficiency virus (HIV) infection inhibits primary strains of HIV type 1 in the presence of natural-killer effector cells. *J. Virol.* **75**, 6953–6961 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Pensiero, A. Fauci, E. Borducchi, S. Clark, R. Hamel, S. King, P. Kozlowski, A. La Porte, G. Landucci, A. Oza, J. Perry, L. Peter, A. Riggs, G. Shaw, N. Simmons, K. Smith, K. Stanley, F. Stephens, Y.-H. Sun, G. Weverling and E. Zablowsky for advice, assistance and reagents. We thank L. Picker, G. Silvestri and B. Walker for critically reviewing this manuscript. The SIV_{MAC239} peptide pools were obtained from the NIH AIDS Research and Reference Reagent Program. We acknowledge support from the US Military Research and Material Command and the US Military HIV Research Program (W81XWH-07-2-0067); the Ragon Institute of MGH, MIT and Harvard; the National Institutes of Health (AI066924, AI078526, AI084794, AI095985, AI060354, AI002642, RR000168); and the Bill and Melinda Gates Foundation. The opinions in this manuscript are those of the authors and do not reflect the views of the US Department of Defense.

Author Contributions D.H.B., M.G.P., H.S., J.C.S., J.G., M.L.R., J.H.K., M.A.M. and N.L.M. designed the study and analysed data. J.L., H.L., D.M.L., M.J.I. and A.S. performed the cellular immunogenicity assays. H.L., M.J.I., M.S.S., G.F., D.N.F., E.A.B. and M.R. performed the humoral immunogenicity assays. L.F.M., P.A., I.O. and V.M.H. prepared the vaccine constructs. A.C. and K.G.M. led the clinical care of the rhesus monkeys. D.S. led the statistical analyses. D.H.B. led the study and wrote the paper with all co-authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.H.B. (dbarouch@bidmc.harvard.edu).

METHODS

Animals, immunizations and challenges. Indian-origin, outbred, young adult, male and female, specific pathogen-free (SPF) rhesus monkeys (*Macaca mulatta*, 80 animals) that did not express the class I alleles *Mamu-A*01*, *Mamu-B*08* and *Mamu-B*17* associated with spontaneous virological control^{11–13} were housed at the New England Primate Research Center (NEPRC), Southborough, Massachusetts, USA. A total of 40 animals was used for each study. Groups were balanced for susceptible and resistant *TRIM5 α* alleles^{1,14}. Immunizations were performed by the intramuscular route in the quadriceps muscles with 2×10^{10} viral particles of Ad35 vectors²⁰, 2×10^{10} viral particles of Ad26 vectors²¹, 10^8 p.f.u. of MVA vectors²⁵, or 5 mg of DNA vaccines expressing SIV_{SME543} Gag-Pol and/or Env gp140. Monkeys were primed at week 0 and boosted at week 24, except DNA vaccine priming that was performed at weeks 0, 4 and 8. To evaluate for protective efficacy and immunological correlates, all monkeys were challenged repetitively beginning at week 52 with six intrarectal inoculations of the heterologous virus SIV_{MAC251} using a 1:1,000 dilution (930 TCID₅₀) of our challenge stock⁹. Monkeys were bled weekly for viral loads (Siemens Diagnostics), and the date of infection was defined as the last challenge time point before the first positive SIV RNA level. Animals were followed to determine set point viral loads. All animal studies were approved by the Harvard Medical School Institutional Animal Care and Use Committee (IACUC).

Cellular immune assays. SIV-specific cellular immune responses were assessed by IFN- γ ELISPOT assays¹⁸ and multiparameter intracellular cytokine staining (ICS) assays^{8,15–17} essentially as described. ELISPOT assays used pools of SIV_{SME543} and SIV_{MAC239} Gag, Pol and Env peptides. Analyses of cellular immune breadth used sub-pools of 10 peptides covering each antigen. Peptides were 15 amino acids in length and overlapped by 11 amino acids. Nine-colour ICS assays used monoclonal antibodies (Becton Dickinson) against CD3 (SP34; Alexa700), CD4 (L200; AmCyan), CD8 (SK1; allophycocyanin-cyanine7 (APC-Cy7)), CD28 (L293; peridinin chlorophyll-A-cyanine5.5 (PerCP-Cy5.5)), CD95 (DX2; phycoerythrin (PE)), CD69 (TP1.55.3; phycoerythrin-Texas Red (energy-coupled dye; ECD); Beckman Coulter), IFN- γ (B27; phycoerythrin-cyanine7 (PE-Cy7)), IL-2 (MQ1-17H12; allophycocyanin (APC)) and TNF- α (Mab11; fluorescein isothiocyanate

(FITC)). IFN- γ backgrounds were consistently $<0.01\%$ in peripheral blood mononuclear cells and $<0.05\%$ in colorectal biopsy specimens.

Humoral immune assays. SIV-specific humoral immune responses were assessed by SIV_{MAC251} Env ELISAs²⁷, TZM-bl luciferase-based virus neutralization assays²⁸ against tier 1 SIV_{SME660} (CP3C-P-A8) and SIV_{MAC251} (TCLA) pseudoviruses, TZM-bl virus neutralization assays against tier 2 SIV_{SME660} (CR54-PK-2A5) and SIV_{MAC251} (SIV_{MAC251.30}) pseudoviruses, antibody-dependent cellular cytotoxicity (ADCC) assays²⁹, and antibody-dependent cell-mediated virus inhibition (ADCVI) assays³⁰. V2-binding assays were performed by surface plasmon resonance with a Biacore 2000 or T200 using a 1:50 serum dilution and a cyclic SIV_{SME543} V2 peptide containing an amino-terminal biotin tag (CIKNNSCAG LEQPMIGCKFNMTGLKRDKKIEYNETWYSRDLICEQPANGSESKCY) and immobilized on streptavidin-coated CM5 chips. Mucosal antibodies were assessed using rectal secretions collected with Weck-Cel sponges. Approximately 100 μ l rectal secretions were eluted and diluted sixfold, and total IgG and IgA as well as SIV Env-specific IgG and IgA (Immune Technology Corporation) were measured by ELISA using a biotin-conjugated anti-monkey IgG and IgA (Alpha Diagnostics) secondary antibody. Mucosal titres were defined as endpoint ELISA titres multiplied by the dilution of the eluted secretions. Samples showed comparable levels of total IgG.

Statistical analyses and immunological correlates. Protection against acquisition of infection was analysed using Wald tests with a proportional hazard model and the exact conditional likelihood method for breaking ties. A discrete time model provided similar estimates. The number of challenges required for 50% infection of each group, hazard ratios with 95% confidence intervals, per-exposure vaccine efficacy and per-exposure risks of infection were quantified. Vaccine efficacy was defined as the reduction in the per-exposure probability of infection as previously described^{4,5}. Analyses of virological and immunological data were performed by Wilcoxon rank-sum tests and analysis of survival by log-rank tests. For these tests, $P < 0.05$ was considered significant and two-tailed tests were performed. Immunological correlates were evaluated by a focused primary analysis and a detailed exploratory analysis using Spearman rank-correlation tests. In the primary analysis, $P < 0.01$ was considered significant, whereas in the exploratory analysis, $P < 0.0014$ was considered significant to adjust for multiple comparisons.

Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos

Michael D. Nodine^{1,2} & David P. Bartel^{1,2}

In animals, maternal gene products deposited into eggs regulate embryonic development before activation of the zygotic genome¹. In plants, an analogous period of prolonged maternal control over embryogenesis is thought to occur based on some gene-expression studies^{2–6}. However, other gene-expression studies and genetic analyses show that some transcripts must derive from the early zygotic genome^{7–14}, implying that the prevailing model does not fully explain the nature of zygotic genome activation in plants. To determine the maternal, paternal and zygotic contributions to the early embryonic transcriptome, we sequenced the transcripts of hybrid embryos from crosses between two polymorphic inbred lines of *Arabidopsis thaliana* and used single-nucleotide polymorphisms diagnostic of each parental line to quantify parental contributions. Although some transcripts seemed to be either inherited from primarily one parent or transcribed from imprinted loci, the vast majority of transcripts were produced in near-equal amounts from both maternal and paternal alleles, even during the initial stages of embryogenesis. Results of reporter experiments and analyses of transcripts from genes that are not expressed in sperm and egg indicate early and widespread zygotic transcription. Thus, in contrast to early animal embryogenesis, early plant embryogenesis is mostly under zygotic control.

The prevailing model for the maternal-to-zygotic transition in plants proposes that most early embryonic messenger RNAs are maternally derived transcripts, resulting either from maternal inheritance or from higher transcriptional activity of maternally derived genes until the globular stages (in which the embryo proper has between ~32 to >100 cells)^{2,6,15}. Because fundamental patterning events, including apical–basal and radial axis formation^{16,17}, occur during the preglobular stages, this model implies that key cell-specification decisions are mostly under maternal control. However, this model is difficult to reconcile with other studies that report equivalent maternal and paternal expression of interrogated genes in preglobular stages^{8,9,11} and zygotic-recessive behaviour of mutants with preglobular developmental phenotypes^{12–14}.

To determine the origins of embryonic transcripts globally, we crossed polymorphic Col-0 and Cvi-0 *Arabidopsis thaliana* accessions and performed RNA-seq on poly(A)⁺ RNA isolated from hybrid embryos with either one-to-two, eight, or ~32 cells in the embryo proper (hereafter referred to as 1-cell/2-cell, 8-cell and ~32-cell embryos). To control for inherent expression differences between Col-0 and Cvi-0 loci, the same procedure was performed using embryos derived from reciprocal crosses. Illumina sequencing of the six samples yielded 73,955,956 reads that both perfectly and uniquely matched the transcribed regions of 23,874 genes (Supplementary Table 1). Overall, transcript levels from the same stage but different reciprocal crosses were highly correlated ($r \geq 0.96$; Supplementary Fig. 1). This reproducibility and sequencing depth indicated that our results would be informative for inferring the maternal, paternal and zygotic contributions to the early embryonic transcriptome.

The prevailing model for the maternal-to-zygotic transition predicted that at the early embryonic stages transcripts would derive primarily

from maternal alleles, and then at the globular stage they would derive more evenly from both alleles because at this stage the zygotic genome would be active. In contrast to this expectation, we found equal amounts of paternally and maternally derived reads at all three stages, including the 1-cell/2-cell and 8-cell embryos (Fig. 1a). When examining transcripts with at least five reads overlapping single-nucleotide polymorphisms (SNPs) in each cross, most were expressed equally from the maternal and paternal alleles, even at the earliest stage (Fig. 1b and Supplementary Data 1; note that in the Col-0 × Cvi-0

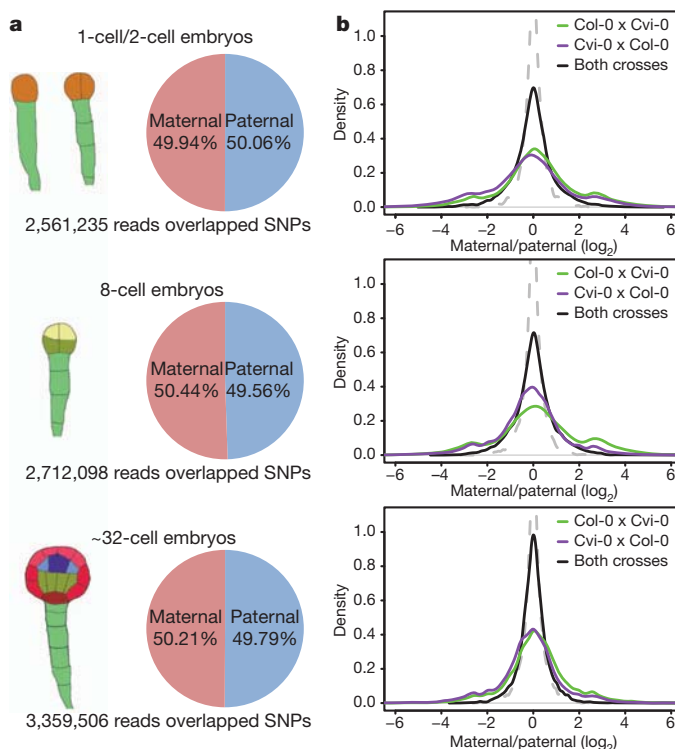


Figure 1 | Maternal and paternal genomes contribute equally to the early embryonic transcriptome. **a**, Proportion of RNA-seq reads overlapping maternal and paternal SNPs within transcribed regions of annotated genes. Drawings illustrate stages of isolated embryos, coloured to represent the diversity of cell lineages in each stage. Pie charts show proportions of reads overlapping maternal and paternal SNPs. **b**, Distributions of maternal-to-paternal ratios (\log_2) in 1-cell/2-cell (top), 8-cell (middle) and ~32-cell (bottom) embryos. Maternal-to-paternal ratios were calculated for transcripts that had ≥ 5 SNP-overlapping reads from the Col-0 × Cvi-0 cross (green; $n > 7,388$ for each stage) or the Cvi-0 × Col-0 cross (purple; $n > 8,071$ for each stage), and transcripts that had ≥ 5 SNP-overlapping reads from both reciprocal crosses (black; $n > 6,892$ for each stage). Binomial modelling determined the distributions expected for transcripts examined in both reciprocal crosses if there had been no parent-of-origin effects and no noise apart from stochastic counting variability (grey dashed lines).

¹Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. ²Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

cross, the maternal parent is Col-0, as the convention is to write the maternal parent first). In each cross, small but discernable subpopulations of transcripts were derived predominantly from either maternal or paternal alleles. However, in the respective reciprocal cross those same transcripts tended to derive from the opposite parent, which showed that rather than arising from parent-of-origin effects, these subpopulations arose mostly from genotypic effects, that is, from preferential expression from either the Col-0 or Cvi-0 alleles. Thus, when considering results of both crosses together, no overall maternal (or paternal) bias was observed, and at each stage the distribution of maternal-to-paternal ratios resembled that predicted for equal contribution from both alleles, with most of the variability explained by stochastic counting statistics, as modelled by the binomial distribution (Fig. 1b). The remainder of the variability was attributed to both additional noise (both experimental and biological) and a small subset of transcripts with parent-of-origin effects, for which the number that were maternally enriched approximately equalled the number that were paternally enriched.

Because most transcripts were derived from the maternal and paternal genomes in near-equal amounts, we hypothesized that either equal amounts of each transcript were inherited upon fertilization or the zygotic genome was activated much earlier than previously proposed. Supporting the latter possibility, the fold-change distributions centred at 0.0 (\log_2), showing no hint of the maternal bias that would be expected if the egg, with its larger cytoplasm, contributed more RNA than the sperm (Fig. 1b). In addition, 543 transcripts had at least fourfold higher reads per million (genome- and cDNA-matching reads) values at the 8-cell stage than at the 1-cell/2-cell stages, which suggested active transcription of the corresponding genes between the 1-cell/2-cell and 8-cell stages (as the alternative model of differential mRNA stability would require a large decrease in total mRNA between these two stages; Supplementary Data 1). Furthermore, 1,138 transcripts that were previously called undetectable in both egg and sperm microarray data sets^{18,19} were among the top 50% most abundant transcripts in our 1-cell/2-cell data sets (Supplementary Data 1). These results, taken together with the observation that transcripts for RNA polymerase II subunits were among the most abundant in the early embryo (Supplementary Table 2), suggested that during the initial stages of embryogenesis many transcripts are transcribed from both maternal and paternal alleles, and that this transcription, combined with turnover of inherited transcripts, quickly overwrites most parent-of-origin biases present when the egg and sperm first fuse.

To test directly whether maternally and paternally inherited genes are transcribed in very early embryos, we used the LhG4/pOp transactivation system²⁰. In this system, one parent contained a transgene encoding the LhG4 transcription factor under the control of either the *RPS5A* or the *UBI3* promoter (chosen because they generate products that can be found in the early embryo^{9,21}), whereas the other parent harboured a nuclear-localized green fluorescent protein (GFP) reporter transgene under the control of the artificial pOp promoter to which LhG4 binds and transcriptionally activates (*pOp::GFP*). Because one gamete contributed the reporter gene and the other gamete contributed its activator, any reporter expression dependent on the activator could not occur until after the zygotic genome was transcriptionally active. No GFP signal was observed in embryos carrying the *pOp::GFP* reporter gene but no activator, which confirmed the absence of leaky expression in either the gametes or embryo (Fig. 2). GFP signal was detected in zygotes within 4–8 h after fertilization from crosses that brought the reporter together with its activator, regardless of whether the reporter was inherited through the egg or sperm (Fig. 2). The GFP signal in the early zygote was not as strong as that in the endosperm, which might indicate a slight delay in activation of the genome of the very early zygote or might result from more robust transcription in the endosperm. Nonetheless, when considered together our results demonstrate that both maternally and paternally inherited chromosomes are transcriptionally active at least in 1-cell embryos and most likely before.

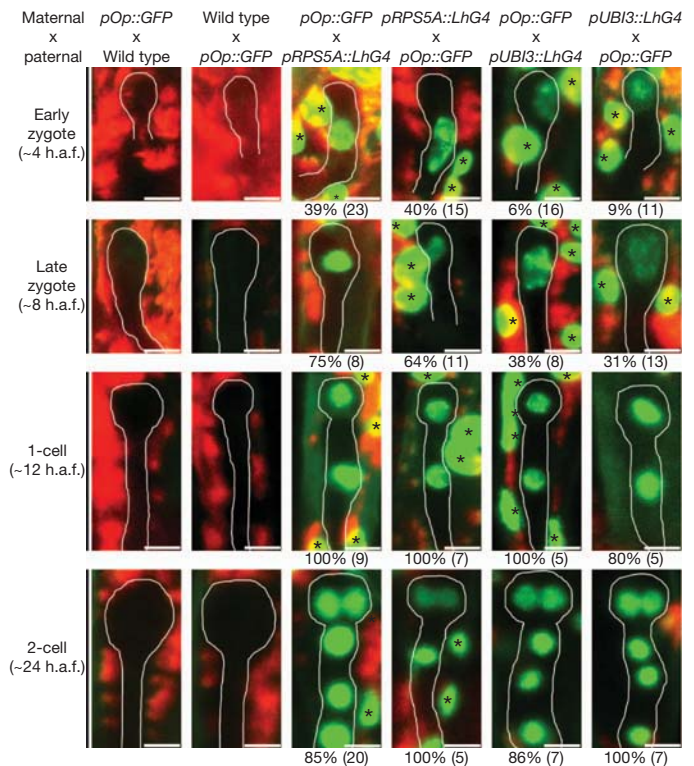


Figure 2 | Maternally and paternally inherited transgenes are transcribed in the initial stages of embryogenesis. Confocal scanning laser microscopy images show early embryos from the indicated crosses, which were between plants with a transgenic reporter (*pOp::GFP*) and either plants without an activator (wild type) or plants with a transgenic activator driven from either the *RPS5A* or the *UBI3* promoter (*pRPS5A::LhG4* or *pUBI3::LhG4*, respectively). Embryonic stages and the approximate number of hours after fertilization (h.a.f.) are shown to the left. GFP signal is indicated by green nuclei; background autofluorescence is shown in red. Embryos are outlined in white for clarity. Asterisks indicate endosperm nuclei. Of the seeds with GFP signal in endosperm (*n* in parentheses), the percentage that also had GFP signal in embryos is indicated below each image. In endosperm, expression from maternal alleles is expected to be twofold higher than that from paternal alleles, but this difference is not observed in these images because the GFP signal in the endosperm was beyond its dynamic range. Scale bars represent 10 μ m.

Having established that most early embryonic transcripts derived equally from the maternal and paternal genomes, we turned our attention to the few that might be preferentially inherited or preferentially expressed from the maternal or paternal alleles. For most early embryonic transcripts the genotype of the allele had a larger effect on transcript levels than did parent of origin (Figs 1b and 3a). Indeed, hundreds of transcripts were preferentially expressed from either the Col-0 or Cvi-0 allele irrespective of their parent of origin (Supplementary Data 2), presumably as direct consequences of either DNA-sequence or epigenetic differences between the Col-0 and Cvi-0 alleles. This prevalence of genotypic/epigenotypic effects over parent-of-origin effects led to significant negative correlations between the maternal-to-paternal ratios at each stage (Fig. 3a, Pearson's $r < -0.43$, $P < 10^{-15}$). Nonetheless, at each stage a small subset of transcripts passed our cutoffs for classification as maternally or paternally enriched, that is, fourfold maternally or paternally enriched in each cross, and not exceeding the false discovery rate (FDR) threshold of 0.05 (Pearson's χ^2 tests, Benjamini and Hochberg FDR). With these cutoffs, 77 and 45 transcripts were designated maternally and paternally enriched, respectively, in at least one stage (Fig. 3a and Supplementary Data 3). For five transcripts, an independent assay involving diagnostic cleavage of an amplified polymorphic sequence was performed, and for all five, the parent-of-origin effects were confirmed (Supplementary Fig. 2).

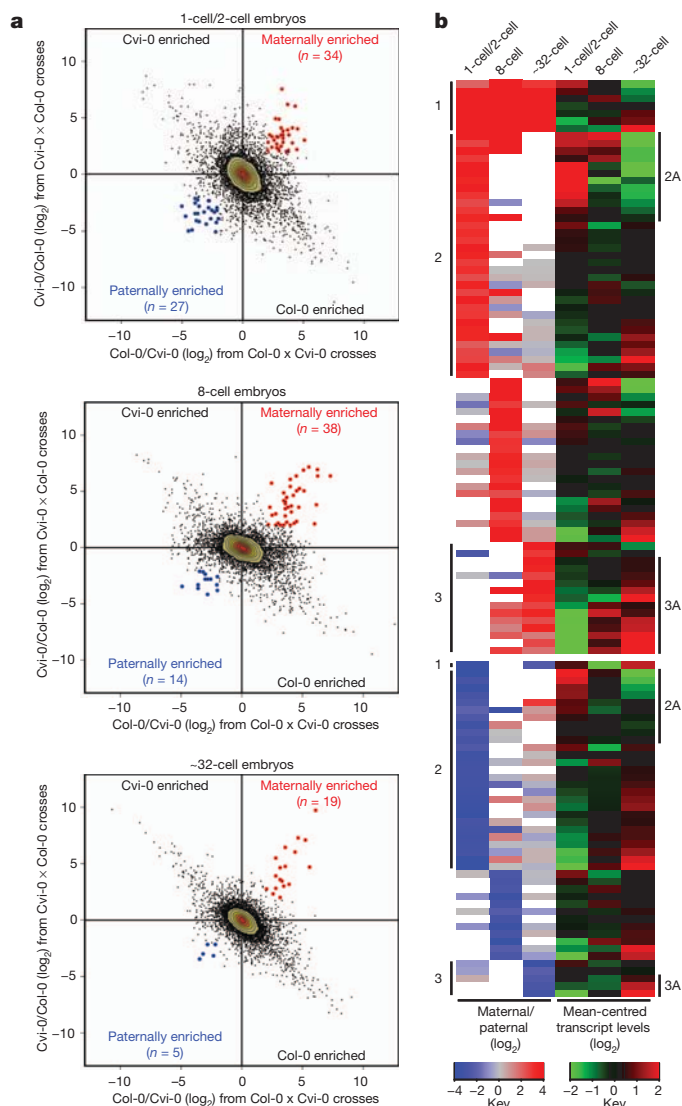


Figure 3 | Dozens of transcripts are preferentially inherited or preferentially expressed. **a**, Maternal-to-paternal ratios for transcripts in 1-cell/2-cell, 8-cell and ~32-cell embryos. Ratios for transcripts that had ≥ 8 SNP-overlapping reads in both the Col-0 \times Cvi-0 and Cvi-0 \times Col-0 crosses are plotted. Points, as well as the numbers, that passed cutoffs for maternally or paternally enriched transcripts (see text) are indicated in red and blue, respectively. **b**, Heat maps of both maternal-to-paternal ratios and changes in transcript levels for transcripts passing cutoffs as maternally (top) or paternally (bottom) enriched in at least one of the three stages. Maternal-to-paternal ratios are colour-coded (key), unless the transcript had < 8 SNP-overlapping reads in one cross, in which case the corresponding heat-map value is not coloured. Changes in the levels of a transcript (reads per million) relative to its mean are colour-coded (key). Numbers to the left indicate candidates for both preferential inheritance and preferential expression (1), preferential inheritance (2) or preferential expression (3). Numbers to the right highlight a subset of the class 2 transcripts, which had relatively high levels in 1-cell/2-cell embryos compared to ~32-cell embryos and thus potentially derived mainly from preferential inheritance (2A), or a subset of class 3 transcripts, which had relatively low levels in 1-cell/2-cell embryos compared to ~32-cell embryos and thus potentially derived mainly from parent-of-origin-specific transcription (3A).

Although the maternal enrichment of some transcripts might be due to their transport from maternal sporophytic tissues, such a mechanism cannot explain paternal enrichment. Some transcripts had strong parent-of-origin effects in 1-cell/2-cell embryos, suggesting that they were preferentially inherited from one parent (Fig. 3b). Confirmation that these biases are indeed due to inheritance would substantially add to the single previously documented example of an

inherited transcript in plants²². Other transcripts had stronger parent-of-origin effects at later stages, suggesting that they were preferentially expressed from either the maternally or the paternally inherited allele (Fig. 3b). This potential parent-of-origin-specific expression implied a form of imprinting in the *Arabidopsis* embryo. Genome-wide approaches similar to ours but looking much later after fertilization greatly expanded the list of genes with parent-of-origin-specific expression in the endosperm but did not identify such genes in embryos^{23,24}. Thus, the imprinting-like phenomenon that we observed in early embryos is short lived, which suggests that it differs from the more persistent imprinting previously characterized in either endosperm or mammals²⁵. For example, it might involve alternative chromatin states inherited from the egg and sperm, which after several cell divisions equilibrate between the two alleles. Such a mechanism would not necessarily require DNA methylation, although we note that in the only previous report of imprinting in plant embryos (at the *MEE1* locus of maize), methylation marks are lost from maternal alleles at the initial stages of embryogenesis, but then re-established to match the paternal alleles at later stages²⁶.

Our results showing that during the initial stages of *Arabidopsis* embryogenesis both the maternal and paternal genomes are active and make essentially equivalent contributions to the embryonic transcriptome are in marked contrast to a recent report that ~88% of the 2-cell/4-cell embryonic transcriptome is derived from the maternal genome². That study, published while our manuscript was in preparation, used SNPs in the transcriptomes of hybrid embryos derived from a cross between the Ler-1 (maternal parent) and Col-0 (paternal parent) accessions to estimate maternal and paternal genomic contributions. Because results from the reciprocal cross were not reported, the previous study could not distinguish parent-of-origin effects from genotypic effects. However, because an analysis unable to account for genotypic effects would be expected to misidentify transcripts as paternally derived as frequently as it misidentifies them as maternally derived, we considered other possibilities for the discrepancy between their results and ours. Our pilot studies had shown that early embryos must be extensively washed to prevent seed-coat RNA contamination, and indeed we found evidence that the embryo RNA samples of ref. 2 contained large amounts of seed-coat mRNA (Supplementary Fig. 3). Because the seed coat is a maternal tissue, this contamination explained why they observed such a large bias in maternal RNAs.

Genes zygotically required for the initial zygotic division have been identified in *Arabidopsis*^{9,12,13}. Moreover, paternal gene expression has been detected for 24 endogenous genes and a transgene in maize zygotes^{8,27}, and *de novo* transcription has been demonstrated recently in tobacco zygotes²⁸. Rather than interpreting these findings as exceptions to the model or as differences between species, our transcriptome-wide analyses and reporter data strongly support the proposal of an alternative model for the maternal-to-zygotic (or, to put it more precisely, the maternal/paternal-to-zygotic) transition in plants. In this model, the zygotic genome is activated within the first few hours after fertilization, and equal transcription of both maternal and paternal alleles generates most of the early embryonic transcriptome at the 1-cell stage or before. Our model implies that although the maternal and paternal gene products are inherited in the zygote upon fertilization and contribute to its development²², the zygotic gene products are the ones that primarily control the initial cell-specification events of embryogenesis.

METHODS SUMMARY

Col-0 and Cvi-0 seed stocks were obtained from the *Arabidopsis* Biological Resource Center. Flowers were emasculated 1 day before crossing, and pools of 20 1-cell/2-cell, 8-cell and ~32-cell embryos were hand-dissected approximately 40, 64 and 78 h after pollination, respectively. RNA isolation, linear amplification of poly(A) RNA and strand-specific RNA-seq were as described²¹, except that embryos were extensively washed before RNA isolation. After removing adaptor sequences, reads were mapped to both the *A. thaliana* genome and transcript models (Col-0 genome reference, TAIR10) with the Bowtie short-read aligner²⁹.

Reads were also mapped to a 'pseudo' Cvi-0 genome and Cvi-0 transcript models, in which SNPs in the Col-0 genome and transcript models were replaced with Cvi-0 variants (ftp://ftp.arabidopsis.org/home/tair/Sequences/Polymorphism_datasets/Ecker_Cvi_snps.txt) generated by the 1001 Genomes Project. Reads that both perfectly and uniquely matched either the genomes or transcript models were retained, and those overlapping transcribed regions of annotated genes were evaluated for overlap with SNPs. Those diagnostic of one genome or the other were tallied after normalizing for differences in sequencing yield between samples. Tallies for multiple SNPs within the same gene were combined before calculating maternal-to-paternal ratios for each transcript.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 September; accepted 1 December 2011.

Published online 22 January 2012.

- Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042 (2009).
- Autran, D. *et al.* Maternal epigenetic pathways control parental contributions to *Arabidopsis* early embryogenesis. *Cell* **145**, 707–719 (2011).
- Vielle-Calzada, J. P., Baskar, R. & Grossniklaus, U. Delayed activation of the paternal genome during seed development. *Nature* **404**, 91–94 (2000).
- Baroux, C., Blanvillain, R. & Gallois, P. Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in *Arabidopsis*. *FEBS Lett.* **509**, 11–16 (2001).
- Grimanelli, D., Perotti, E., Ramirez, J. & Leblanc, O. Timing of the maternal-to-zygotic transition during early seed development in maize. *Plant Cell* **17**, 1061–1072 (2005).
- Baroux, C., Autran, D., Gillmor, C. S., Grimanelli, D. & Grossniklaus, U. The maternal to zygotic transition in animals and plants. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 89–100 (2008).
- Springer, P. S., Holding, D. R., Groover, A., Yordan, C. & Martienssen, R. A. The essential Mcm7 protein PROLIFERA is localized to the nucleus of dividing cells during the G₁ phase and is required maternally for early *Arabidopsis* development. *Development* **127**, 1815–1822 (2000).
- Meyer, S. & Scholten, S. Equivalent parental contribution to early plant zygotic development. *Curr. Biol.* **17**, 1686–1691 (2007).
- Weijers, D., Geldner, N., Offringa, R. & Jurgens, G. Seed development: Early paternal gene activity in *Arabidopsis*. *Nature* **414**, 709–710 (2001).
- Sprunck, S., Baumann, U., Edwards, K., Langridge, P. & Dresselhaus, T. The transcript composition of egg cells changes significantly following fertilization in wheat (*Triticum aestivum* L.). *Plant J.* **41**, 660–672 (2005).
- Ingouff, M. *et al.* Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in *Arabidopsis*. *Curr. Biol.* **20**, 2137–2143 (2010).
- Xu, J. *et al.* EMBRYONIC FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo transition in *Arabidopsis*. *Plant J.* **42**, 743–758 (2005).
- Ronceret, A. *et al.* The first zygotic division in *Arabidopsis* requires *de novo* transcription of thymidylate kinase. *Plant J.* **53**, 776–789 (2008).
- Muralla, R., Lloyd, J. & Meinke, D. Molecular foundations of reproductive lethality in *Arabidopsis thaliana*. *PLoS ONE* **6**, e28398 (2011).
- Pillot, M. *et al.* Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* **22**, 307–320 (2010).
- Mayer, U., Torres Ruiz, R. A., Berleth, T., Misera, S. & Jurgens, G. Mutations affecting body organization in the *Arabidopsis* embryo. *Nature* **353**, 402–407 (1991).
- Laux, T., Wurschum, T. & Breuninger, H. Genetic regulation of embryonic pattern formation. *Plant Cell* **16** (Suppl. 1), S190–S202 (2004).
- Wuest, S. E. *et al.* *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Curr. Biol.* **20**, 506–512 (2010).
- Borges, F. *et al.* Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiol.* **148**, 1168–1181 (2008).
- Moore, I., Galweiler, L., Grosskopf, D., Schell, J. & Palme, K. A transcription activation system for regulated gene expression in transgenic plants. *Proc. Natl Acad. Sci. USA* **95**, 376–381 (1998).
- Nodine, M. D. & Bartel, D. P. MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev.* **24**, 2678–2692 (2010).
- Bayer, M. *et al.* Paternal control of embryonic patterning in *Arabidopsis thaliana*. *Science* **323**, 1485–1488 (2009).
- Gehring, M., Missirian, V. & Henikoff, S. Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. *PLoS ONE* **6**, e23687 (2011).
- Hsieh, T. F. *et al.* Regulation of imprinted gene expression in *Arabidopsis* endosperm. *Proc. Natl Acad. Sci. USA* **108**, 1755–1762 (2011).
- Feil, R. & Berger, F. Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet.* **23**, 192–199 (2007).
- Jahnke, S. & Scholten, S. Epigenetic resetting of a gene imprinted in plant embryos. *Curr. Biol.* **19**, 1677–1681 (2009).
- Scholten, S., Lorz, H. & Kranz, E. Paternal mRNA and protein synthesis coincides with male chromatin decondensation in maize zygotes. *Plant J.* **32**, 221–231 (2002).
- Zhao, J. *et al.* Dynamic changes of transcript profiles after fertilization are associated with *de novo* transcription and maternal elimination in tobacco zygote, and mark the onset of the maternal-to-zygotic transition. *Plant J.* **65**, 131–145 (2011).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Ecker and the 1001 Genomes Project for generating the list of Cvi-0 SNPs; the J. Harada and R. Goldberg laboratories for making the *Arabidopsis* seed development LCM microarray data sets publicly available; I. Moore for the transactivation vectors; J. Long for the *UBI3* promoter fragment; the Whitehead Genome Technology Core for sequencing; and D. Meinke for providing a curated list of preglobular zygotic-recessive mutants before publication. This work used the W. M. Keck Biological Imaging Facility at the Whitehead Institute and was supported by NIH grant GM067031 (D.P.B.) and NIH Postdoctoral Fellowship GM084656 (M.D.N.). D.P.B. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions M.D.N. designed and performed the experiments. M.D.N. and D.P.B. interpreted the results and wrote the manuscript.

Author Information Raw and processed RNA-seq data sets have been deposited into NCBI GEO (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE33713. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.P.B. (dbartel@wi.mit.edu).

METHODS

Plant growth and embryo dissections. Col-0 and Cvi-0 seed stocks were obtained from the *Arabidopsis* Biological Resource Center. Plants were grown at 22 °C in a Conviron growth chamber with a 16-h light/8-h dark cycle. Flowers were emasculated 1 day before crossing, and pools of 20 1-cell/2-cell, 8-cell and ~32-cell embryos were hand-dissected approximately 40, 64 and 78 h after pollination, respectively.

Embryo dissections were performed in water on glass slides using a SZH stereo microscope (Olympus) with 128× magnification. Tungsten needles (Fine Science Tools) were used to cut open the seeds and remove the embryos from the rest of the seed material. For each embryo, a heat-pulled borosilicate glass microcapillary pipette (Harvard Apparatus) was used to move it to a depression slide containing 50 µl water. The embryo was then washed two additional times by serial transfer to slides containing 50 µl water, using a new pipette and fresh water for each transfer. Washed embryos were collected in 30 µl of RNAlater (Ambion).

RNA-seq and analyses. RNA isolation, linear amplification of poly(A) RNA and strand-specific RNA-seq were as described²¹, except that embryos were extensively washed before RNA isolation. After removing adaptor sequences, reads were mapped to both the *A. thaliana* genome and transcript models (Col-0 genome reference, TAIR10) with the Bowtie short-read aligner²⁹. Reads were also mapped to a 'pseudo' Cvi-0 genome and Cvi-0 transcript models, in which SNPs in the Col-0 genome and transcript models were replaced with Cvi-0 variants (ftp://ftp.arabidopsis.org/home/tair/Sequences/Polymorphism_datasets/Ecker_Cvi_snps.txt) generated by the 1001 Genomes Project. Reads that both perfectly and uniquely matched either the genomes or transcript models were retained, and those overlapping transcribed regions of annotated genes were evaluated for overlap with SNPs. To normalize for differences in library sizes, the numbers of reads representing each transcript were divided by the total number of reads matching the genome and transcript models. SNP-overlapping reads were assigned to one of the parental genomes and tallied for each transcript, combining tallies for multiple SNPs within the same transcript.

Statistical analyses and graphics. Statistical analyses were performed and associated graphics were generated with the R statistical computing base package³⁰. The SciPy tools for Python were used to calculate χ^2 test statistics and associated probabilities.

Transgenic lines. The pBIN+LhG4-GW vector was generated by digesting both the pBIN+LhG4 vector and the attR1/attR2-containing vector GWRFa::pET42a

with KpnI and ligating together the appropriate digestion fragments. *RPS5A* and *UBI3* promoter fragments were amplified with appropriate primer pairs, cloned into pENTR-D/TOPO (Invitrogen) and then recombined with pBIN+LhG4-GW to generate pRPS5A::LhG4-GW and pUBI3::LhG4-GW, respectively. The pV-TOP-GFP vector was generated by double-digesting pV-TOP(E3) and NLSGFPF1R1::pCR8 with HindIII and PmeI, and ligating together the appropriate digestion fragments to replace the GUS reporter with a nuclear-localized GFP reporter. NLSGFPF1R1::pCR8 was created by amplifying pCGTAG with appropriate primers and cloning into pCR8/GW-TOPO (Invitrogen). pOp::GFP was then generated by digesting pV-TOP-GFP and GWF1R3::pENTR with SalI, and ligating together the appropriate digestion fragments. GWF1R3::pENTR was created by amplifying the attR1/attR2 Gateway cassette from GWRFa::pET42a with a series of overlap extension PCRs to remove internal SalI sites and add SalI sites on the ends of amplicons, which were then cloned into pENTR/D-TOPO (Invitrogen). Oligonucleotides used to generate the above constructs are listed in Supplementary Table 3. All constructs were transformed into Col-0 by *Agrobacterium*-mediated transformation. Because the progeny of T₁ lines gave more robust GFP signal than did progeny of established lines, T₁ lines were used for all reporter crosses. Embryos from crosses using at least 13 *pOp::GFP*, four *pRPS5A::LhG4* and six *pUBI3::LhG4* independent T₁ lines were examined.

Microscopy. For confocal scanning laser microscopy, developing seeds were mounted in 50 mM potassium phosphate buffer, pH 7.2, with 5% glycerol. A 488-nm laser on a Zeiss LSM 510 confocal microscope was used to excite GFP and seed autofluorescence, and images were collected at 505–530 nm and 644–719 nm, respectively.

Analyses of cleaved amplified polymorphic sequences. Transcripts containing a SNP that created or disrupted a restriction site in the corresponding cDNA were selected for analysis. Random hexamers (Invitrogen) were used for reverse transcription with Superscript III (Invitrogen) and primers flanking the SNPs were used for amplification (Supplementary Table 3). The amplified DNA was digested with the appropriate restriction enzyme (New England Biolabs; Supplementary Table 4). Digestion products were resolved on 2% agarose gels stained with ethidium bromide, and bands were quantified with Quantity One 1-D analysis software (Bio-Rad).

30. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Austria, 2010).

Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities

Lynn M. Boyden¹, Murim Choi^{1*}, Keith A. Choate^{2*}, Carol J. Nelson-Williams¹, Anita Farhi¹, Hakan R. Toka³, Irina R. Tikhonova⁴, Robert Bjornson⁴, Shrikant M. Mane⁴, Giacomo Colussi⁵, Marcel Lebel⁶, Richard D. Gordon⁷, Ben A. Semmekrot⁸, Alain Poujol⁹, Matti J. Välimäki¹⁰, Maria E. De Ferrari⁵, Sami A. Sanjad¹¹, Michael Gutkin¹², Fiona E. Karet¹³, Joseph R. Tucci¹⁴, Jim R. Stockigt¹⁵, Kim M. Keppler-Noreuil¹⁶, Craig C. Porter¹⁷, Sudhir K. Anand¹⁸, Margo L. Whiteford¹⁹, Ira D. Davis²⁰, Stephanie B. Dewar²¹, Alberto Bettinelli²², Jeffrey J. Fadrowski²³, Craig W. Belsha²⁴, Tracy E. Hunley²⁵, Raoul D. Nelson²⁶, Howard Trachtman²⁷, Trevor R. P. Cole²⁸, Maury Pinsk²⁹, Detlef Bockenhauer³⁰, Mohan Shenoy³¹, Priya Vaidyanathan³², John W. Foreman³³, Majid Rasoulpour³⁴, Farook Thameem³⁵, Hania Z. Al-Shahrouri³⁵, Jai Radhakrishnan³⁶, Ali G. Gharavi³⁶, Beatrice Goilav³⁷ & Richard P. Lifton¹

Hypertension affects one billion people and is a principal reversible risk factor for cardiovascular disease. Pseudohypoaldosteronism type II (PHAII), a rare Mendelian syndrome featuring hypertension, hyperkalaemia and metabolic acidosis, has revealed previously unrecognized physiology orchestrating the balance between renal salt reabsorption and K^+ and H^+ excretion¹. Here we used exome sequencing to identify mutations in kelch-like 3 (*KLHL3*) or cullin 3 (*CUL3*) in PHAII patients from 41 unrelated families. *KLHL3* mutations are either recessive or dominant, whereas *CUL3* mutations are dominant and predominantly *de novo*. *CUL3* and BTB-domain-containing kelch proteins such as *KLHL3* are components of cullin-RING E3 ligase complexes that ubiquitinate substrates bound to kelch propeller domains^{2–8}. Dominant *KLHL3* mutations are clustered in short segments within the kelch propeller and BTB domains implicated in substrate⁹ and cullin⁵ binding, respectively. Diverse *CUL3* mutations all result in skipping of exon 9, producing an in-frame deletion. Because dominant *KLHL3* and *CUL3* mutations both phenocopy recessive loss-of-function *KLHL3* mutations, they may abrogate ubiquitination of *KLHL3* substrates. Disease features are reversed by thiazide diuretics, which inhibit the Na-Cl cotransporter in the distal nephron of the kidney; *KLHL3* and *CUL3* are expressed in this location, suggesting a mechanistic link between *KLHL3* and *CUL3* mutations, increased Na-Cl reabsorption, and disease pathogenesis. These findings demonstrate the utility of exome sequencing in disease gene identification despite the combined complexities of locus heterogeneity, mixed models of transmission and frequent *de novo* mutation, and establish a fundamental role for *KLHL3* and *CUL3* in blood pressure, K^+ and pH homeostasis.

A small number of genes causing Mendelian forms of hypertension have been identified, establishing the role of increased renal salt reabsorption in its pathogenesis^{10–12}. The study of PHAII has identified a physiological mechanism that orchestrates activities of diverse electrolyte flux pathways, allowing maximal salt reabsorption in response to aldosterone when angiotensin II (AII) is elevated, as in settings of reduced intravascular volume (hypovolaemia), versus maximal potassium secretion in settings of hyperkalaemia, in which aldosterone is elevated without changes in AII¹. The role of WNK kinases in this process was revealed by discovery of their mutation in a small fraction of PHAII kindreds¹¹. Dominant gain-of-function mutations in *WNK4* or *WNK1* lead to constitutively increased salt reabsorption in the distal nephron regardless of volume status, resulting in hypertension, and inhibition of K^+ secretion despite marked hyperkalaemia^{1,11,13–17}.

We studied a cohort of 52 PHAII kindreds, including 126 affected subjects with renal hyperkalaemia and otherwise normal renal function; hypertension and acidosis were present in 71% and 82%, respectively. There was wide variation in disease severity and age of clinical presentation (Supplementary Figs 1 and 2). Mutations in *WNK1* or *WNK4* were present in only seven of these kindreds (13%). Those without *WNK* mutations had only 2.0 ± 1.4 affected members, complicating mapping efforts.

Exome sequencing of eleven unrelated PHAII index cases without *WNK* mutations was performed. Index cases and affected relatives (five trios and one quartet) were also subjected to genome-wide SNP genotyping. Tabulation of high quality novel protein-altering variants revealed 124 genes with three or more variants, 50 with four or more, and 23 with five. Concurrent analysis of linkage among the multiplex families was used to prioritize loci harbouring variants that co-segregated with

¹Department of Genetics and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ²Department of Dermatology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. ³Renal Division, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁴Yale Center for Genome Analysis, Yale University, New Haven, Connecticut 06510, USA. ⁵Nephrology Unit, Niguarda-Ca' Granda Hospital, Milan 20162, Italy. ⁶Department of Medicine, Laval University, Québec G1K 7P4, Canada. ⁷Endocrine Hypertension Research Centre, University of Queensland School of Medicine, Brisbane QLD 4006, Australia. ⁸Department of Pediatrics, Canisius Wilhelmina Hospital, Nijmegen 6500 GS, Netherlands. ⁹Department of Pediatrics, Pays d'Aix Hospital, Aix-en-Provence 13616, France. ¹⁰Division of Endocrinology, Department of Medicine, Helsinki University Central Hospital, Helsinki 00290, Finland. ¹¹Department of Pediatrics and Adolescent Medicine, American University Medical Center, Beirut 1107 2020, Lebanon. ¹²Hypertension Research Center, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 07103, USA. ¹³Department of Medical Genetics, University of Cambridge, Cambridge CB2 1TN, United Kingdom. ¹⁴Division of Endocrinology, Roger Williams Medical Center, Providence, Rhode Island 02908, USA. ¹⁵Department of Endocrinology and Diabetes, Ewen Downie Metabolic Unit, Alfred Hospital, Melbourne VIC 3004, Australia. ¹⁶Division of Medical Genetics, Department of Pediatrics, University of Iowa Children's Hospital, Iowa City, Iowa 52242, USA. ¹⁷Division of Nephrology, Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA. ¹⁸Department of Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA. ¹⁹Duncan Guthrie Institute of Medical Genetics, Royal Hospital for Sick Children, Glasgow G3 8SJ, United Kingdom. ²⁰Baxter Healthcare Corporation, McGaw Park, Illinois 60085, USA. ²¹Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA. ²²Division of Pediatrics, Mandic Hospital, Merate 23807, Italy. ²³Department of Pediatrics, Johns Hopkins School of Medicine, Baltimore, Maryland 21287, USA. ²⁴Division of Nephrology, Department of Pediatrics, Saint Louis University Health Sciences Center, St Louis, Missouri 63110, USA. ²⁵Division of Nephrology, Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA. ²⁶Division of Nephrology, Department of Pediatrics, University of Utah, Salt Lake City, Utah 84132, USA. ²⁷Division of Nephrology, Cohen Children's Medical Center of New York, New Hyde Park, New York 11040, USA. ²⁸West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham B15 2TG, United Kingdom. ²⁹Division of Nephrology, Department of Pediatrics, University of Alberta, Edmonton T6G 2M7, Canada. ³⁰Renal Unit, University College London Institute of Child Health, London WC1N 1EH, United Kingdom. ³¹Department of Nephrology, Royal Manchester Children's Hospital, Manchester M27 4HA, United Kingdom. ³²Department of Endocrinology, Children's National Medical Center, Washington, District of Columbia 20010, USA. ³³Department of Pediatrics, Duke University Medical Center, Durham, North Carolina 27710, USA. ³⁴Division of Nephrology, Connecticut Children's Medical Center, Hartford, Connecticut 06106, USA. ³⁵Division of Nephrology, Department of Medicine, University of Texas Health Science Center, San Antonio, Texas 78229, USA. ³⁶Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA. ³⁷Division of Nephrology, Children's Hospital at Montefiore, Bronx, New York 10467, USA.

*These authors contributed equally to this work.

disease; this identified 28 genes with novel protein-altering variants that co-segregated with disease in two or more multiplex families. This revealed *KLHL3* as a strong candidate, with novel *KLHL3* mutations comprising five alleles in three kindreds, all of which co-segregated with the trait. These include one kindred in which affected members are homozygous for a nonsense mutation (W470X), one in which affected members are compound heterozygotes for two missense mutations (F322C and S410L), and one segregating a heterozygous missense mutation (R528H). As a confirmation of significance, Fisher's exact test was used to compare the prevalence of novel protein-altering variants in all genes in PHAII cases versus 699 control exomes. A single gene, *KLHL3*, showed a burden of mutation that surpassed genome-wide significance ($P = 1.1 \times 10^{-8}$; Supplementary Tables 1–3).

KLHL3 was sequenced in all PHAII index cases, identifying novel mutations in 24 (Fig. 1a, b and Supplementary Figs 3 and 4). Nearly all are at positions conserved among orthologues (Supplementary Fig. 5). Sixteen kindreds have heterozygous mutations that co-segregate with the trait under a dominant model (log of odds (lod) score 6.9, < -2 under other models). In contrast, eight index cases inherited mutations in both *KLHL3* alleles. In these kindreds, affected members are confined to siblings of index cases who inherited the same two mutations, whereas unaffected relatives inherited zero or one mutation (lod score 4.3 for a recessive model, < -2 for other models). Recessive transmission has not been previously described for PHAII. Consistent with two modes of transmission, subjects with dominant *KLHL3* mutations had significantly higher serum K^+ levels (6.2 ± 0.6 mM)

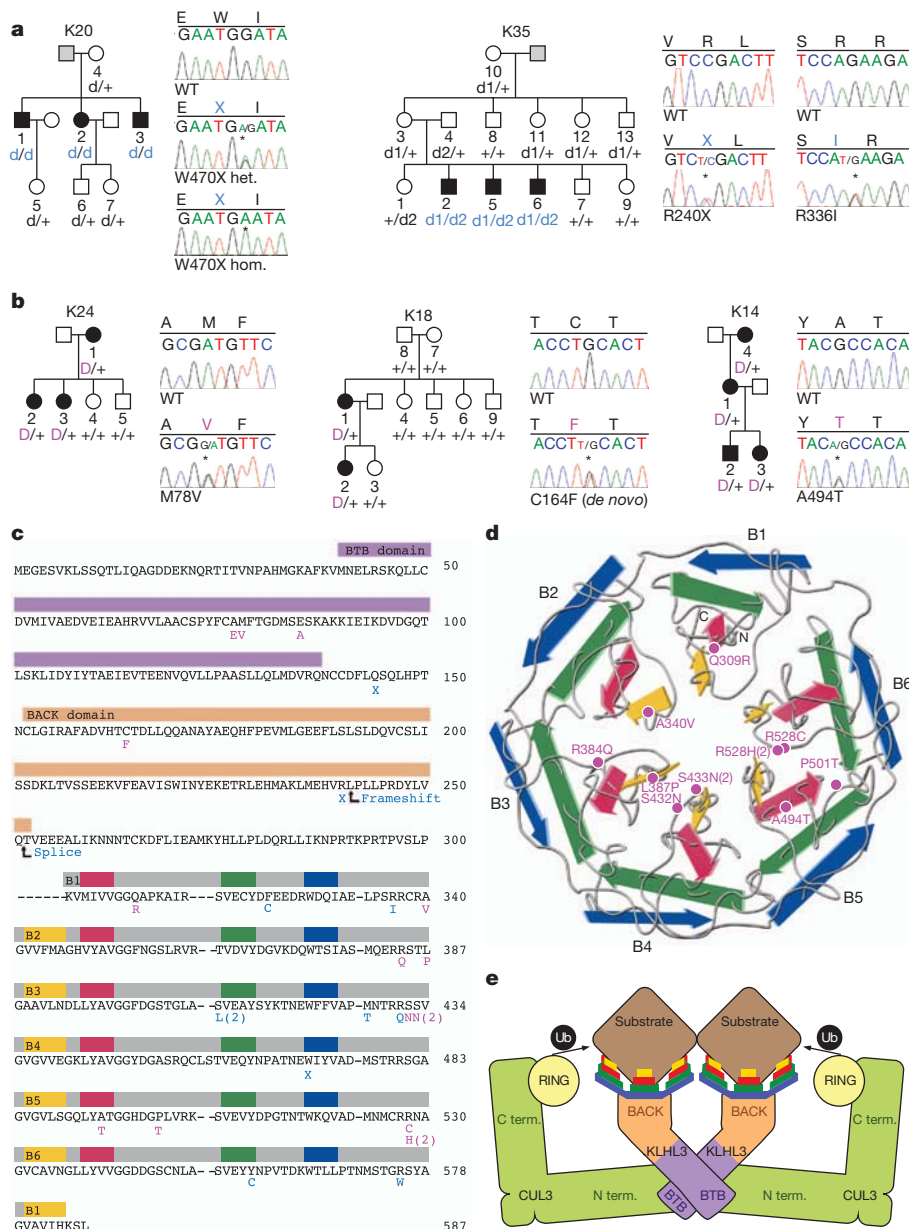


Figure 1 | Recessive and dominant *KLHL3* mutations in PHAII kindreds. **a, b**, Representative kindreds demonstrating recessive (**a**) and dominant (**b**) *KLHL3* mutations (all 24 kindreds are shown in Supplementary Figs 3 and 4). Affected, unaffected and phenotype-undetermined subjects are denoted by black, white and grey symbols, respectively. *KLHL3* alleles are denoted by '+' (wild type), 'd' (recessive mutation) and 'D' (dominant mutation). Sequence traces show wild-type (WT) and mutant (*) alleles and encoded amino acids. het., heterozygous; hom., homozygous. **c**, *KLHL3* protein sequence. Coloured bars indicate BTB domain (lavender), BACK domain (peach) and kelch

propeller blades (B1–B6, grey) with β -strands 'a–d' in yellow, red, green and blue, respectively. Recessive (light blue) and dominant (pink) mutations are shown; recurrences are indicated by numbers. **d**, Kelch propeller schematic, from the *KLHL2* crystal structure²⁷. β -Strands are coloured as in **c**; dominant mutations are indicated. **e**, CRL schematic, comprising a BTB-containing kelch protein (*KLHL3*), CUL3 and a ubiquitin (Ub)-transfer-mediating RING protein, with substrate bound via the kelch propeller. The complex is shown as a dimer⁷. C term., C terminus; N term., N terminus.

than heterozygotes for recessive mutations (4.8 ± 0.6 mM) ($P < 10^{-4}$, Student's *t*-test; normal range 3.5–5.0 mM). These findings establish that PHAII can be caused by either recessive or dominant *KLHL3* mutations. Importantly, we infer that mutations in dominant kindreds are probably dominant-negative, because they phenocopy the features of recessive disease.

KLHL3 contains an amino-terminal BTB domain, a BACK domain, and carboxy-terminal kelch-like repeats that form a six-bladed β -propeller structure^{2,4,5} (Fig. 1c–e). There are over 50 BTB-containing kelch (BTB-kelch) genes in humans⁴; their propeller domains bind substrate proteins, promoting substrate ubiquitination via interaction of the BTB domain with CUL3, a component of a cullin–RING E3 ubiquitin ligase (CRL)^{3,5,6}. Ubiquitination serves diverse functions, including targeting proteins for proteasomal degradation as well as non-degradative roles such as modulation of protein activity, interaction and localization^{7,8}.

Whereas recessive *KLHL3* mutations are distributed throughout the encoded protein, dominant *KLHL3* mutations show marked clustering (Fig. 1c). Nine of sixteen dominant mutations alter one of the last four amino acids of the six 'd–a' loops that connect the outermost ('d') β -strand of one kelch propeller blade to the innermost ('a') β -strand of the next blade. Two others are in 'b–c' loops. These dominant PHAII mutations lie near the hub of the propeller (Fig. 1d) at or near sites implicated in substrate binding in paralogues⁹ (Supplementary Fig. 5). Three other dominant mutations cluster within the BTB domain, at or

near sites implicated in cullin binding in paralogues⁵. We infer that dominant mutations in *KLHL3* probably impair binding either to specific substrates or to CUL3.

After accounting for *KLHL3*, *WNK1* and *WNK4* mutations, 21 PHAII kindreds without mutations remained. We considered the presumed functional partner of *KLHL3*, *CUL3* (Fig. 1e), as a potential candidate. Among PHAII exomes, novel heterozygous *CUL3* variants were suggested in two. Sequencing *CUL3* in all index cases identified 17 with novel heterozygous mutations, all in cases without *KLHL3*, *WNK1* or *WNK4* mutations (Fig. 2a and Supplementary Fig. 6). Eight of these mutations were documented to be *de novo*, providing overwhelming evidence that these mutations are disease-causing. *CUL3* mutations all cluster in sites implicated in splicing of exon 9, including the intron 8 splice acceptor ($n = 4$), the intron 9 splice donor ($n = 5$), the putative intron 8 splice branch site ($n = 5$), and a putative splice enhancer in exon 9 ($n = 3$, within a TTGGA(T/A)) splice enhancer consensus sequence¹⁸) (Fig. 2b).

To test the impact of these mutations on splicing, *CUL3* genomic DNA spanning exon 8 to exon 10 containing either wild-type sequence or one of nine PHAII mutations was cloned and expressed in HEK293 cells, and the spliced RNA products were analysed. Whereas the wild-type sequence produces a properly spliced product containing all three exons, each of the mutants produces a predominant product that skips exon 9, joining exon 8 to exon 10 (Fig. 2c, d). This results in an in-frame 57 amino acid deletion (residues 403–459) in the segment linking the

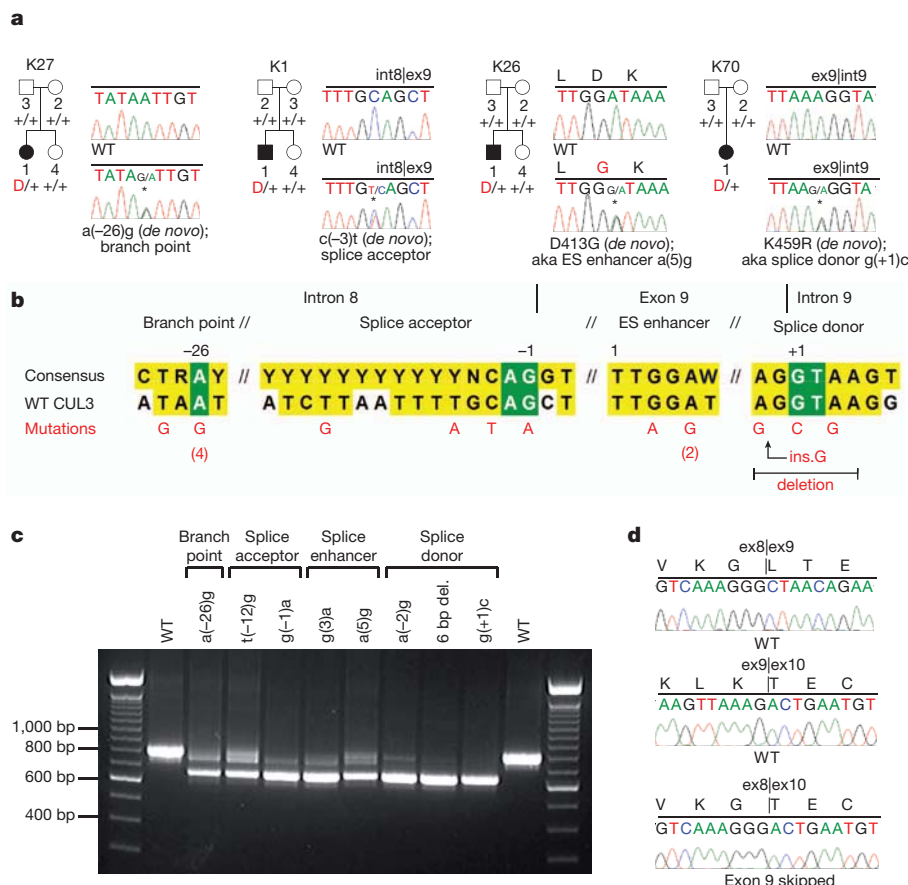


Figure 2 | Dominant *CUL3* mutations in PHAII kindreds cause skipping of exon 9. **a**, Representative kindreds demonstrating *CUL3* mutations, depicted as in Fig. 1 (all 17 kindreds are shown in Supplementary Fig. 6). int8, intron 8; ex9, exon 9; int9, intron 9. Positions are numbered relative to splice sites and the first base of the exonic splice (ES) enhancer. **b**, *CUL3* mutation locations. Consensus splicing sequences^{18,28} and corresponding wild-type *CUL3* sequences are shown; invariant bases (green) and consensus homology (yellow) are indicated. Mutations are shown in red; recurrences are indicated by

numbers. Ins.G, insertion of a G (guanine). **c**, Reverse transcription–polymerase chain reaction (RT–PCR) of spliced RNA. Wild-type *CUL3* constructs produce a single product including exons 8, 9 and 10 (844 bp); all nine mutants tested produce a predominant product that skips exon 9 (673 bp). **d**, Representative RT–PCR sequences. The wild-type construct produces complementary DNA with properly spliced junctions between exons 8–9 (top) and 9–10 (middle), whereas the mutant construct (splice donor g(+1)c) produces cDNA joining exon 8 to exon 10 (bottom).

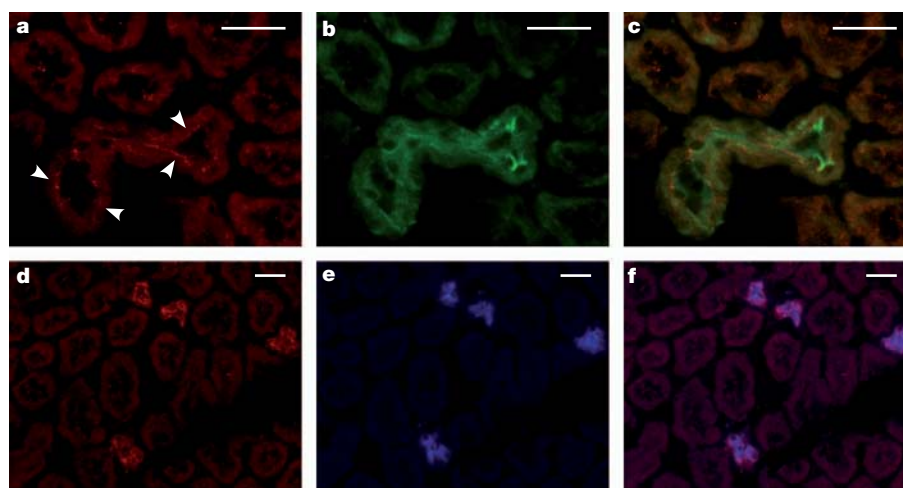


Figure 3 | *KLHL3* expression in the kidney. **a–f**, Mouse kidney sections stained with antibodies to *KLHL3* (red), *TRPM6* (a marker of the DCT²⁹, green) and *AQP2* (a marker of the collecting duct³⁰, blue). Scale bars, 25 μ m. **a–c**, Staining for *KLHL3* (**a**), *TRPM6* (**b**) and the merged image

(**c**) demonstrates *KLHL3* expression in the DCT with apical localization (arrowheads). **d–f**, Staining for *KLHL3* (**d**), *AQP2* (**e**) and the merged image (**f**) demonstrates *KLHL3* expression in the collecting duct.

BTB-binding and RING-binding domains of *CUL3*. The fact that *CUL3* mutations phenocopy recessive *KLHL3* mutations suggests that they abrogate *CUL3* function at *KLHL3* targets.

As with PHAII caused by *WNK1* and *WNK4* mutations¹, virtually all patients with *KLHL3* and *CUL3* mutations have been treated with thiazide diuretics, inhibitors of the Na–Cl cotransporter (NCC), resulting in correction of phenotypic abnormalities. *WNK4* regulates the activities of NCC^{13,14,16}, the epithelial Na⁺ channel ENaC¹⁷, and the K⁺ channel ROMK¹⁵, and is co-expressed with these proteins in the renal distal convoluted tubule (DCT) and collecting duct^{11,19,20}. Staining mouse kidney sections with specific antibodies demonstrated that *KLHL3* is predominantly expressed in the DCT and collecting duct, with apical localization in the DCT (Fig. 3). *CUL3* is ubiquitously expressed and is in all nephron segments, with particularly high expression in the proximal tubule, but also in the DCT and collecting duct (Supplementary Fig. 7). These findings are consistent with both proteins having a role in the regulation of salt and electrolyte homeostasis in the distal nephron.

There are highly significant differences in phenotypic severity among PHAII patients with mutations in different genes (Table 1 and Supplementary Figs 1, 2 and Supplementary Table 4). Subjects with *CUL3* mutations presented at much younger ages than those with mutation in *KLHL3*, *WNK1* or *WNK4*, had significantly more severe hyperkalaemia and metabolic acidosis, and were far more likely to have hypertension before age 18 (others commonly develop hypertension at later ages). The majority of subjects with *CUL3* mutations demonstrated failure to thrive or growth impairment. These observations, in conjunction with the high rate of *CUL3* *de novo* mutation, support impairment of reproductive fitness. Among the other mutant loci, there remain significant differences in disease severity (rank order recessive *KLHL3* > dominant *KLHL3* > *WNK4* > *WNK1*).

KLHL3 and *CUL3* mutations account for 79% of kindreds in our cohort. Gene identification was complicated by the combined effects of locus heterogeneity, two modes of transmission at one locus, and few informative meioses. Many so far unsolved Mendelian traits may have similar complexities. Use of control exomes as comparators for analysis of mutation burden may be broadly applicable to discovery of such loci.

The most parsimonious mechanism of *KLHL3* and *CUL3* mutations is that they abrogate ubiquitination of targets normally bound by *KLHL3*, activity that is required for normal modulation of renal salt, K⁺ and H⁺ handling in response to physiological challenge; this speculation will require biochemical verification. The fact that recessive mutations in *KLHL3* cause PHAII without other diverse effects implies either that *KLHL3* targets are highly restricted to the renal salt and electrolyte pathway, or that loss of *KLHL3* function at other targets can be compensated by other loci. BTB-kelch–*CUL3* CRLs can act as dimers, with two substrate-binding domains capable of engaging the same target molecule⁷. This suggests a potential mechanism to explain dominant-negative effects of *KLHL3* and *CUL3* mutations.

CUL3-based CRLs participate in a wide range of critical cellular processes⁸ via binding diverse BTB-domain-containing proteins^{3,5}. *CUL3* mutations affecting all or many of these activities would undoubtedly produce very broad phenotypes. *CUL3* mutations in PHAII merely phenocopy the effects of loss of *KLHL3*, suggesting that they selectively abrogate function at *KLHL3* targets. The stereotypic consequences of *CUL3* mutations, all deleting 57 amino acids in a region linking the BTB-binding and RING-binding domains, support such a specific effect. Consistent with this possibility, introduction of a flexible linker sequence to this region of *CUL1* leaves substrate protein binding and ubiquitin polymerization intact, but nonetheless abolishes ubiquitination of a normal substrate²¹.

Table 1 | PHAII phenotypes, stratified by genotype.

Mutant gene	No. of kindreds	No. of affecteds	Dx/Ref age*	K ⁺ (mM) (nl 3.5–5.0 mM)†	HCO ₃ [−] (mM) (nl 22–28 mM)†	Hypertension at ≤age 18 (%)†
<i>CUL3</i>	17	21	9 ± 6	7.5 ± 0.9	15.5 ± 2.0	94
<i>KLHL3</i> recessive	8	14	26 ± 14	6.8 ± 0.5	17.6 ± 1.5	14
<i>KLHL3</i> dominant	16	40	24 ± 18	6.2 ± 0.6	17.2 ± 2.5	17
<i>WNK4</i>	5	15	28 ± 18	6.4 ± 0.7	20.8 ± 2.3	10
<i>WNK1</i>	2	23	36 ± 20	5.8 ± 0.8	22.4 ± 4.6	13

Dx/Ref age, age at diagnosis or referral. K⁺, serum potassium; HCO₃[−], serum bicarbonate; nl, normal range. Hypertension at ≤age 18 (%), percentage of affected patients diagnosed with hypertension by age 18. Values for Dx/Ref age, K⁺ and HCO₃[−] are means ± standard deviations. Significance of differences among genotype classes was calculated by ANOVA (Dx/Ref age, K⁺, HCO₃[−]) or Fisher's exact test (hypertension).

**P* = 0.0002.

†*P* < 0.0001.

Thiazide diuretics correct abnormalities in virtually all PHAII subjects; similar correction is seen in a mouse model of PHAII with either thiazides or genetic ablation of NCC¹⁶. These findings suggest that increased NCC activity is likely to be a common pathogenic mechanism. Co-expression of KLHL3 and CUL3 with NCC in the DCT, and evidence that NCC is ubiquitinated²² is consistent with this notion. ROMK and the H⁺ ATPase are respectively required for net renal K⁺ secretion²³ and H⁺ secretion²⁴ and are also likely targets, although their activity is expected to be decreased, rather than increased, in PHAII. Another thiazide-sensitive Na–Cl cotransport pathway in the collecting duct has recently been described, suggesting an additional potential target²⁵. Whether a KLHL3–CUL3 CRL acts directly or indirectly on these targets, whether they alter delivery of NCC and other targets to, or retrieval from, the plasma membrane, and what upstream pathways regulate this activity is unknown. Similarly, whether KLHL3–CUL3 and WNKs operate within the same or different pathways is presently unknown; it is of interest that segments of NRF2 that interact with the kelch propeller domain of KEAP1 are highly acidic^{9,26}, akin to the domain of WNK4 that is mutated in PHAII¹¹.

These findings demonstrate previously unrecognized roles for KLHL3 and CUL3. Understanding the upstream regulators and downstream targets of KLHL3–CUL3 activity will provide further insight into the mechanisms underlying maintenance of blood pressure and electrolyte homeostasis in response to diverse environmental challenges.

METHODS SUMMARY

A cohort of 52 PHAII kindreds comprising 126 affected subjects was ascertained, characterized and recruited for study. Index cases of eleven kindreds were chosen for whole exome capture and sequencing, and novel variants were identified. Genes were prioritized for follow-up and subjected to Sanger sequencing of index cases of each kindred. Segregation of rare variants within kindreds was analysed. KLHL3 and CUL3 were localized in kidney by staining with specific antibodies and immunofluorescence microscopy. Effects of CUL3 mutations on RNA splicing were studied by analysis of spliced products produced in mammalian cells.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 September; accepted 22 December 2011.

Published online 22 January 2012.

- Kahle, K. T., Ring, A. M. & Lifton, R. P. Molecular physiology of the WNK kinases. *Annu. Rev. Physiol.* **70**, 329–355 (2008).
- Lai, F. *et al.* Molecular characterization of KLHL3, a human homologue of the *Drosophila kelch* gene. *Genomics* **66**, 65–75 (2000).
- Furukawa, M., He, Y. J., Borchers, C. & Xiong, Y. Targeting of protein ubiquitination by BTB–Cullin 3–Roc1 ubiquitin ligases. *Nature Cell Biol.* **5**, 1001–1007 (2003).
- Prag, S. & Adams, J. C. Molecular phylogeny of the kelch-repeat superfamily reveals an expansion of BTB/kelch proteins in animals. *BMC Bioinformatics* **4**, 42 (2003).
- Stogios, P. J., Downs, G. S., Jauhal, J. J., Nandra, S. K. & Prive, G. G. Sequence and structural analysis of BTB domain proteins. *Genome Biol.* **6**, R82 (2005).
- Hudson, A. M. & Cooley, L. Phylogenetic, structural and functional relationships between WD- and kelch-repeat proteins. *Subcell. Biochem.* **48**, 6–19 (2008).
- Zimmerman, E. S., Schulman, B. A. & Zheng, N. Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* **20**, 714–721 (2010).
- Sarikas, A., Hartmann, T. & Pan, Z. Q. The cullin protein family. *Genome Biol.* **12**, 220 (2011).
- Lo, S. C., Li, X., Henzl, M. T., Beamer, L. J. & Hannink, M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J.* **25**, 3605–3617 (2006).
- Lifton, R. P., Gharavi, A. G. & Geller, D. S. Molecular mechanisms of human hypertension. *Cell* **104**, 545–556 (2001).

- Wilson, F. H. *et al.* Human hypertension caused by mutations in WNK kinases. *Science* **293**, 1107–1112 (2001).
- Choi, M. *et al.* K⁺ channel mutations in adrenal aldosterone-producing adenomas and hereditary hypertension. *Science* **331**, 768–772 (2011).
- Wilson, F. H. *et al.* Molecular pathogenesis of inherited hypertension with hyperkalemia: the Na–Cl cotransporter is inhibited by wild-type but not mutant WNK4. *Proc. Natl Acad. Sci. USA* **100**, 680–684 (2003).
- Yang, C. L., Angell, J., Mitchell, R. & Ellison, D. H. WNK kinases regulate thiazide-sensitive Na–Cl cotransport. *J. Clin. Invest.* **111**, 1039–1045 (2003).
- Kahle, K. T. *et al.* WNK4 regulates the balance between renal NaCl reabsorption and K⁺ secretion. *Nature Genet.* **35**, 372–376 (2003).
- Lalioti, M. D. *et al.* Wnk4 controls blood pressure and potassium homeostasis via regulation of mass and activity of the distal convoluted tubule. *Nature Genet.* **38**, 1124–1132 (2006).
- Ring, A. M. *et al.* WNK4 regulates activity of the epithelial Na⁺ channel *in vitro* and *in vivo*. *Proc. Natl Acad. Sci. USA* **104**, 4020–4024 (2007).
- Fairbrother, W. G., Yeh, R. F., Sharp, P. A. & Burge, C. B. Predictive identification of exonic splicing enhancers in human genes. *Science* **297**, 1007–1013 (2002).
- Bachmann, S., Bostanjoglo, M., Schmitt, R. & Ellison, D. H. Sodium transport-related proteins in the mammalian distal nephron—distribution, ontogeny and functional aspects. *Anat. Embryol. (Berl.)* **200**, 447–468 (1999).
- Welling, P. A. & Ho, K. A comprehensive guide to the ROMK potassium channel: form and function in health and disease. *Am. J. Physiol. Renal Physiol.* **297**, F849–F863 (2009).
- Zheng, N. *et al.* Structure of the Cul1–Rbx1–Skp1–F box^{Skp2} SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).
- Ko, B. *et al.* RasGRP1 stimulation enhances ubiquitination and endocytosis of the sodium–chloride cotransporter. *Am. J. Physiol. Renal Physiol.* **299**, F300–F309 (2010).
- Simon, D. B. *et al.* Genetic heterogeneity of Bartter's syndrome revealed by mutations in the K⁺ channel, ROMK. *Nature Genet.* **14**, 152–156 (1996).
- Karet, F. E. *et al.* Mutations in the gene encoding B1 subunit of H⁺-ATPase cause renal tubular acidosis with sensorineural deafness. *Nature Genet.* **21**, 84–90 (1999).
- Leviel, F. *et al.* The Na⁺-dependent chloride-bicarbonate exchanger SLC4A8 mediates an electroneutral Na⁺ reabsorption process in the renal cortical collecting ducts of mice. *J. Clin. Invest.* **120**, 1627–1635 (2010).
- Tong, K. I. *et al.* Different electrostatic potentials define ETGE and DLG motifs as hinge and latch in oxidative stress response. *Mol. Cell. Biol.* **27**, 7511–7521 (2007).
- Wang, Y. *et al.* MMDB: annotating protein sequences with Entrez's 3D-structure database. *Nucleic Acids Res.* **35**, D298–D300 (2007).
- Zhang, M. Q. Statistical features of human exons and their flanking regions. *Hum. Mol. Genet.* **7**, 919–932 (1998).
- Voets, T. *et al.* TRPM6 forms the Mg²⁺ influx channel involved in intestinal and renal Mg²⁺ absorption. *J. Biol. Chem.* **279**, 19–25 (2004).
- Fushimi, K. *et al.* Cloning and expression of apical membrane water channel of rat kidney collecting tubule. *Nature* **361**, 549–552 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the PHAII subjects, their families, and the health care professionals whose participation made this study possible; S. Umlauf and the staff of the Yale Center for Genome Analysis; J. Santosuosso; H. Tirrell and the staff of Beckman Coulter Genomics; V. Klump, Y. Lu, U. Scholl and J. Zhou for providing reagents; W. Hill for artistic assistance with Fig. 1d; and E. Boyden, S. Boyden, L. Cooley and M. Hochstrasser for helpful discussions. This work was supported in part by the Leducq Transatlantic Network on Hypertension and grants from the National Institutes of Health (P30-DK079310 and UL1-RR024139).

Author Contributions L.M.B., M.C., K.A.C. and R.P.L. designed experiments and analysed data. L.M.B., C.J.N.-W. and I.R.T. performed experiments. A.F., H.R.T., G.C., M.L., R.D.G., B.A.S., A.P., M.J.V., M.E.D.F., S.A.S., M.G., F.E.K., J.R.T., J.R.S., K.M.K.-N., C.C.P., S.K.A., M.L.W., I.D.D., S.B.D., A.B., J.J.F., C.W.B., T.E.H., R.D.N., H.T., T.R.P.C., M.P., D.B., M.S., P.V., J.W.F., M.R., F.T., H.Z.A.-S., J.R., A.G.G. and B.G. recruited PHAII subjects and families. R.B. and S.M.M. directed the information technology and DNA sequencing infrastructure. L.M.B. and R.P.L. wrote the manuscript.

Author Information mRNA and protein sequences are available at NCBI under accession numbers NM_017415.2 and NP_059111.2 (KLHL3), NM_003590.3 and NP_003581.1 (CUL3); mutation data is available at dbSNP under batch accession 1056535. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.P.L. (richard.lifton@yale.edu).

METHODS

Study subjects. Index cases were referred for PHAII. Patients and participating family members provided consent to a study protocol approved by the Yale Human Investigation Committee. Control exomes were 699 unrelated subjects of European ancestry without hypertension, sequenced as part of diverse gene discovery projects. Genomic DNA was isolated from venous blood using standard methods.

Exome capture, sequencing and variant calling. Genomic DNA from eleven PHAII index cases and 699 controls was captured on NimbleGen 2.1M human exome arrays (Roche) and sequenced on the Illumina GenomeAnalyzer as previously described³¹. Reads were mapped to the reference genome (hg18) using Maq³² and genotypes of targeted bases were called with SAMtools³³. Variants found in dbSNP v.130 or 1000 Genomes databases were excluded from further analysis. Remaining variants were considered 'novel' and annotated for impact on the encoded protein, conservation and expression³¹. Aligned reads were viewed with the Integrative Genomics Viewer³⁴. Among PHAII cases, 94.2% of targeted bases were read by eight or more independent reads; sensitivity and specificity of heterozygous calls were estimated at 93.7% and 99.9% by comparison to Illumina SNP genotyping. Among controls, 94.4% of targeted bases were read by eight or more independent reads; sensitivity and specificity of heterozygous calls were estimated at 94.5% and 99.8%. Sanger sequencing of 212 novel variants from controls with a SAMtools quality score of ≥ 100 demonstrated validation in 211 and amplification failure in 1, supporting high specificity of variant calls.

SNP genotyping and linkage analysis. For the 11 PHAII index cases and their affected relatives (five trios and one quartet) genome-wide SNP genotyping was performed using Illumina Human610-Quad BeadChips and GenomeStudio software. Approximately 40K tag SNPs were extracted using Plink³⁵. Analysis of linkage was performed using Merlin³⁶, specifying an autosomal dominant model with no phenocopies. Variants from exome sequencing in regions of the genome that were excluded (lod score < -2) were removed from further analysis, while those that supported linkage were prioritized for further evaluation. In kindreds showing potential recessive transmission of PHAII, SNP genotypes were examined for regions of homozygosity, and linkage was performed specifying an autosomal recessive model.

Sanger sequencing of KLHL3 and CUL3. PCR amplification and Sanger sequencing from genomic DNA was performed using standard methods. Primers were designed with Primer3³⁷. Variants identified by exome sequencing were verified. All exons and flanking intronic sequences of *KLHL3* and *CUL3* were sequenced from all PHAII index cases. Previously unidentified mutations were discovered and verified by independent amplification and sequencing. Co-segregation of mutations with disease was determined by sequencing in all available kindred members. *CUL3* exon 9 and its flanking intronic sequence was sequenced in 150 unaffected unrelated controls, none of whom were found to harbour previously unidentified variants. It is noteworthy that because of lower or absent sequence coverage at or near intron-exon junctions, splice donor and acceptor mutations in *CUL3* were suggested in two of the eleven PHAII exomes (SAMtools quality scores 96 and 75) but three branch site mutations were outside the exome sequence and one splice enhancer mutation was poorly covered (SAMtools quality score 3).

Genome-wide assessment of mutation burden. Genes show substantial variation in the prevalence of novel or rare protein-altering variants for biological reasons, including differences in gene size and variation in the proportion of bases that are under purifying selection, and for technical reasons, including difficulties in accurately mapping short sequence reads among closely related paralogs. These factors can limit the ability to directly identify disease loci by simply counting and ranking genes according to the absolute number of such variants, particularly for diseases with substantial locus or model heterogeneity. This gene-to-gene variation can be accounted for by use of control exome data. The prevalence of rare variation in each gene in case exomes was compared to the corresponding prevalence in a large set of control exomes with a Fisher's exact test. Variants included in the analysis were protein-altering (missense, nonsense and splice site mutations) and high quality (≥ 8 independent reads and SAMtools quality score ≥ 100).

For a gene-wise test of rare variant burden in a genome with $\sim 21,000$ genes, correction for multiple testing suggests a threshold *P* value of $\sim 2.4 \times 10^{-6}$, anticipated to produce a false discovery rate (FDR) of one gene per twenty experiments. The FDR of the Fisher's test was evaluated by Monte Carlo simulation, which confirmed an FDR of < 1 gene per 20 experiments (Supplementary Table 1). The power to identify trait-related loci was estimated as a function of the number of variants detected in cases and the number of case exomes sequenced (Supplementary Table 2), and the test was applied comparing the eleven PHAII and 699 control exomes (Supplementary Table 3).

Orthologue and paralogue comparisons. Protein sequences of orthologues and paralogues were aligned with Clustal W³⁸. Crystal structures were examined with Cn3D³⁹. The locations of KLHL3 propeller mutations were compared to the crystal structure of human KLHL2 (PDB accession 2XN4)²⁷, the closest human paralogue⁴ (85% amino acid identity in the propeller). The location of the peptide encoded by *CUL3* exon 9 was approximated by comparison to the crystal structure of human CUL1 (PDB accessions 1LDJ and 1LDK)^{22,27}.

Splicing assay. A 3,782 bp segment of *CUL3*, extending from 287 bp proximal to exon 8 to 327 bp distal to exon 10, was amplified by PCR (Advantage 2 polymerase, Clontech) from genomic DNA of nine PHAII patients with different *CUL3* mutations and one subject with wild-type *CUL3* sequence. Products were cloned into the pcDNA6.2/GW/D-TOPO mammalian expression vector (Invitrogen), and plasmids were purified (QIAprep, Qiagen) and sequenced. HEK293 cells were transfected independently with each plasmid using Lipofectamine 2000 (Invitrogen) and harvested ~ 24 h post-transfection. RNA was isolated using RNeasy with DNase on-column digestion (Qiagen). The spliced expression products were assessed by reverse transcription with oligo(dT) priming (Superscript III RT, Invitrogen) followed by PCR with vector-specific and *CUL3*-specific primers. Products were fractionated and visualized via agarose gel electrophoresis, and sequenced. Untransfected and water controls were negative.

Immunofluorescence. Fresh frozen mouse kidney sections were fixed with ethanol at 4 °C for 30 min and acetone at -20 °C for 3 min, washed with 1 \times PBS, and permeabilized with 0.1% Triton X-100 (Sigma) at room temperature (22 °C) for 10 min. Sections were blocked with 10% donkey serum and 1% bovine serum albumin at room temperature for 1 h, incubated with primary antibodies at room temperature for 1 h or 4 °C overnight, washed four times with 1 \times PBS, incubated with secondary antibody at room temperature for 1 h, and washed four times with 1 \times PBS, with DAPI nuclear counterstain in the second wash. Slides were mounted with Mowiol (Polysciences) and 1% n-propyl gallate (Sigma) as an anti-fade agent. Primary antibodies included 1:100 rabbit anti-KLHL3, 1:50 rabbit anti-CUL3, 1:800 or 1:1,200 guinea pig anti-TRPM6 (ab66655, ab1871, and ab47017; Abcam) and 1:400 or 1:800 goat anti-AQP2 (sc-9882, Santa Cruz Biotechnology). Secondary antibodies, diluted 1:800, included donkey Cy3 anti-rabbit, Cy2 anti-guinea pig and 649 anti-goat IgG (AffiniPure, Jackson ImmunoResearch). Staining with secondary antibodies only was consistently negative.

- Choi, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl Acad. Sci. USA* **106**, 19096–19101 (2009).
- Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**, 1851–1858 (2008).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnol.* **29**, 24–26 (2011).
- Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genet.* **30**, 97–101 (2002).
- Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**, 365–386 (2000).
- Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
- Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. & Bryant, S. H. Cn3D: sequence and structure views for Entrez. *Trends Biochem. Sci.* **25**, 300–302 (2000).

A novel sensor to map auxin response and distribution at high spatio-temporal resolution

Géraldine Brunoud¹, Darren M. Wells^{2*}, Marina Oliva^{1*}, Antoine Larrieu^{2,3*}, Vincent Mirabet¹, Amy H. Burrow⁴, Tom Beeckman³, Stefan Kepinski⁴, Jan Traas¹, Malcolm J. Bennett² & Teva Vernoux¹

Auxin is a key plant morphogenetic signal¹ but tools to analyse dynamically its distribution and signalling during development are still limited. Auxin perception directly triggers the degradation of Aux/IAA repressor proteins^{2–6}. Here we describe a novel Aux/IAA-based auxin signalling sensor termed DII-VENUS that was engineered in the model plant *Arabidopsis thaliana*. The VENUS fast maturing form of yellow fluorescent protein⁷ was fused in-frame to the Aux/IAA auxin-interaction domain (termed domain II; DII)⁵ and expressed under a constitutive promoter. We initially show that DII-VENUS abundance is dependent on auxin, its TIR1/AFBs co-receptors^{4–6,8} and proteasome activities. Next, we demonstrate that DII-VENUS provides a map of relative auxin distribution at cellular resolution in different tissues. DII-VENUS is also rapidly degraded in response to auxin and we used it to visualize dynamic changes in cellular auxin distribution successfully during two developmental responses, the root gravitropic response and lateral organ production at the shoot apex. Our results illustrate the value of developing response input sensors such as DII-VENUS to provide high-resolution spatio-temporal information about hormone distribution and response during plant growth and development.

Central to auxin signalling is the ubiquitin- and proteasome-dependent degradation of Aux/IAA catalysed by the SCF-type E3 ubiquitin-ligase complexes SCF^{TIR1/AFB1–5} (refs 2–6, 8). Aux/IAA

repressors form heterodimers with transcription factors termed auxin response factors (ARFs)^{9,10}. Auxin directly promotes the interaction between TIR1/AFBs auxin co-receptors and Aux/IAAs⁵, thus recruiting Aux/IAAs to the SCF complex^{3,6} and derepressing ARF-bound loci. This allows the transcription of target genes including most *Aux/IAA* genes, hence providing a negative feedback loop (Fig. 1a)^{2,10}. The most widely used tools to monitor auxin distribution in *planta* are DR5-based auxin-inducible reporters whose promoter contains several ARF binding sites^{11,12}. However, as an output of the auxin response pathway (Fig. 1a), reporter activity does not directly relate to endogenous auxin abundance but also reflects the contribution of a complex signalling pathway².

Monitoring the degradation of an Aux/IAA-based green fluorescent protein (GFP) reporter would provide a better target for an auxin sensor as its signal can be related more directly to hormone abundance (Fig. 1a)^{3,5,6,13}. This has proved very challenging because Aux/IAA half-lives are often shorter than GFP maturation time^{7,14–17}. To overcome this technical limitation, we fused the VENUS fast maturing yellow fluorescent protein (YFP)⁷ to the auxin-interaction domain (termed domain II; DII)⁵ from several Aux/IAA proteins and expressed these fusion proteins under the constitutive 35S promoter (Fig. 1b and Supplementary Fig. 1a, b). Confocal imaging of transgenic root apical tissues revealed similar fluorescence patterns but the strongest signal

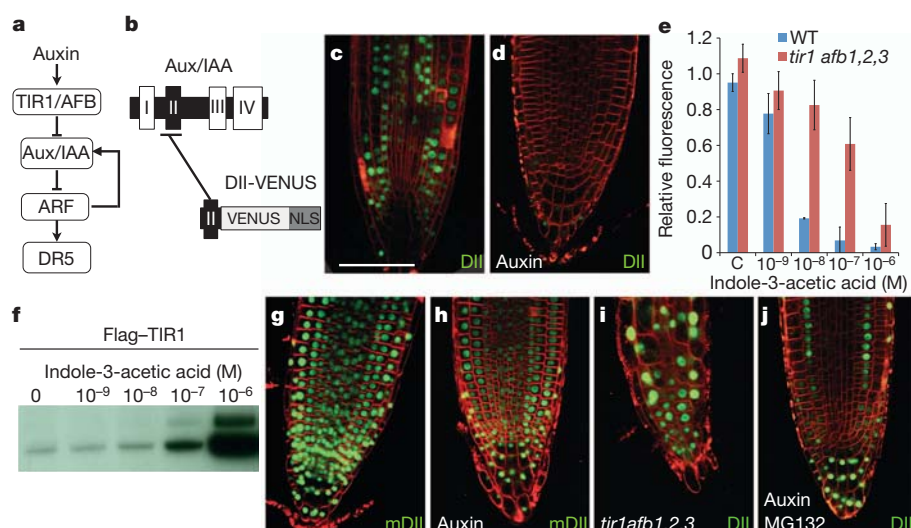


Figure 1 | DII-VENUS degradation is dependent on auxin, TIR1/AFBs and proteasome activity. **a**, Schematic representation of auxin signalling. **b**, Design of DII-VENUS; NLS, nuclear localization signal. **c**, **d**, DII-VENUS fluorescence in absence (**c**) or presence (**d**) of 1 μ M indole-3-acetic acid. **e**, Dose-dependent degradation of DII-VENUS in wild-type and mutant roots. **f**, IAA28 domain II peptide pull-down assay with Flag-TIR1 in the presence of indole-3-acetic acid.

g, **h**, mDII-VENUS fluorescence in the absence (**g**) or presence (**h**) of 1 μ M indole-3-acetic acid. **i**, DII-VENUS signal in co-receptor quadruple mutant. **j**, DII-VENUS signal upon 1 μ M indole-3-acetic acid and proteasome inhibitor (MG132) co-treatment. Green channel, VENUS; red channel, FM4-64. Scale bar, 50 μ m. Error bars, s.d. (WT, $n = 2$; mutant, $n = 3$).

¹Laboratoire de Reproduction et Développement des Plantes, CNRS, INRA, ENS Lyon, UCBL, Université de Lyon, 69364 Lyon, France. ²Centre for Plant Integrative Biology, University of Nottingham, Sutton Bonington LE12 5RD, UK. ³Department of Plant Systems Biology, VIB, Ghent University, B-9052 Ghent, Belgium. ⁴Centre for Plant Science, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK.

*These authors contributed equally to this work.

was obtained using domain II of the most stable Aux/IAA used, IAA28 (Fig. 1c and Supplementary Fig. 1c–f). We thus focused our analyses on this form of the sensor, henceforth called DII-VENUS.

We tested on root tissues the relationship between auxin, its response components and DII-VENUS in several independent ways. First, the DII-VENUS signal was sensitive in a dose-dependent fashion to exogenous auxin treatment (Fig. 1c–e). Second, the DII peptide interacted with its co-receptors TIR1 (Fig. 1f), AFB1 and AFB5 (ref. 18) in an auxin-dependent manner. Third, introducing a mutation in the domain II sequence of DII-VENUS (mDII-VENUS), that disrupts the interaction between Aux/IAA, auxin and the TIR1/AFBs⁵, reduced the differential distribution of fluorescence (Fig. 1g; see below for description of pattern) and blocked its auxin-induced degradation (Fig. 1h). Fourth, DII-VENUS fluorescence was ubiquitously distributed in roots of the most strongly affected *tir1 afb1 afb2 afb3* quadruple mutant⁴ (Fig. 1i) and the mutants were significantly less sensitive to auxin treatment (Fig. 1e). Fifth, disruption of ubiquitin-dependent breakdown of Aux/IAA proteins using proteasome inhibitors stabilized DII-VENUS and blocked its auxin-induced degradation (Fig. 1j and Supplementary Fig. 2). We conclude that DII-VENUS abundance is regulated by auxin via its receptors, consistent with the model for Aux/IAA degradation (Fig. 1a)². We also demonstrated that DII-VENUS does not disrupt the activity of the auxin response machinery (Supplementary Fig. 3 and Supplementary Information). Hence, DII-VENUS directly reports, but does not interfere with, the input into the auxin signalling pathway.

We next took advantage of the simple cellular organization of the root apex to quantify the distribution of DII-VENUS fluorescence with cellular definition (Fig. 2a, b and Supplementary Fig. 4). Because the TIR1/AFB1–3 co-receptor distribution shows only limited variations in the root meristem region (except in the root cap; Supplementary Fig. 5a–d)¹⁹, this will confer a homogeneous perception capacity. Hence, the spatial distribution of DII-VENUS fluorescence is likely to represent an inverted auxin distribution map in the root tip. This conclusion is further supported by the more homogeneous fluorescence distribution of mDII-VENUS (Fig. 2c and Supplementary Fig. 6) and by the complementary patterns of DII-VENUS and *DR5::VENUS* expression in the quiescent centre, columella and differentiating xylem cells (Fig. 2d, e)^{11,20,21}. However, mDII-VENUS fluorescence distribution suggests a higher 35S promoter activity in the epidermis and cortex in the elongation zone and in the most external root cap cells (Fig. 2c and Supplementary Fig. 6). Lower expression of

the TIR/AFBs is also expected to confer a lower sensitivity to auxin in the root cap (Supplementary Fig. 5a–d). In both cases, this will lead to an underestimation of auxin levels by DII-VENUS. Analyses of TIR/AFBs co-receptor distribution and of 35S promoter activity are thus essential to interpret the DII-VENUS pattern.

Even considering these biases (Supplementary Information), DII-VENUS quantification indicates that auxin levels are reproducibly higher in the first two tiers of columella cells and initials, the quiescent centre, the stele initials and early daughters and the differentiating xylem cells (Fig. 2b and Supplementary Fig. 4)²². The other cells in the root meristem have lower levels of auxin, with minima observed in the epidermis and cortex, but auxin levels significantly increased close to the start of the elongation zone. This increase occurs closer to the root tip in the epidermis and vasculature compared to the cortex (Fig. 2b). The DII-VENUS fluorescence map thus confirms the local maximum of auxin at the quiescent centre and in the columella cells¹¹ but also allows visualizing the distribution of auxin in the entire root tip. It also reveals a previously unsuspected auxin accumulation starting at the transition zone between the meristem and the elongation zone. These results are in partial agreement with measurements of auxin concentrations in root tissues obtained after cell sorting²³, the differences being possibly due to the higher resolution achieved using DII-VENUS.

We also detected differential distributions of DII-VENUS fluorescence in the vegetative shoot apical meristem (SAM), in the vascular tissues of the hypocotyl (Fig. 2f, g) and later during development in the inflorescence SAM and young floral meristems (Supplementary Fig. 7a–d)¹⁸. As in the root, reduced differential expression with mDII-VENUS, partly complementary *DR5::VENUS* patterns and distribution of TIR1/AFB1–3 co-receptors (Fig. 2h, i and Supplementary Fig. 5e–h) indicate that the distribution of DII-VENUS fluorescence is primarily controlled by auxin levels in the shoot apex (Supplementary Information)¹⁸. DII-VENUS is therefore able to report relative auxin distribution at high spatial resolution in various tissues and developmental stages. In addition, in both root and shoot tissues, DII-VENUS is degraded not only in cells where *DR5* is expressed but also in cells that do not express *DR5* (Figs 1c, 2a–f and Supplementary Fig. 7 and Supplementary Information)¹⁸. This observation demonstrates that the Aux/IAA-ARF signalling pathway contributes significantly to the definition of the *DR5* expression pattern.

To analyse the temporal resolution of the DII-VENUS sensor, we compared dynamic changes in DII-VENUS and *DR5::VENUS* signals

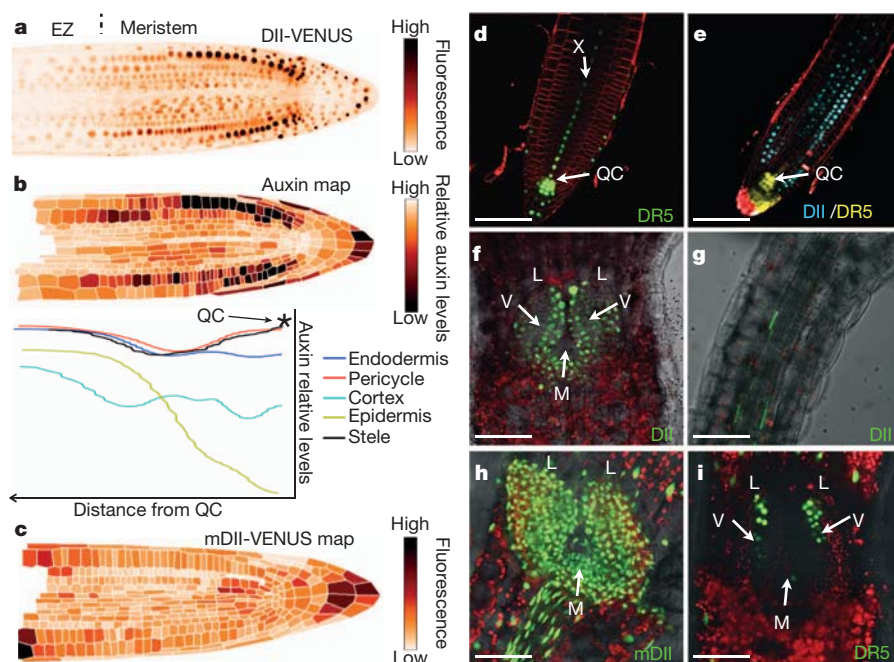


Figure 2 | DII-VENUS provides a sensor to map auxin distribution in plant tissues. a–d, Root meristematic tissues. **a, b,** DII-VENUS fluorescence (**a**) and corresponding map of relative auxin distribution (**b**; top) using the same look-up table. EZ, elongation zone. Tissue-specific changes in auxin levels along the root axis are shown (**b**; bottom). QC, quiescent centre. **c,** Fluorescence map of mDII-VENUS control. **d,** *DR5::VENUS*. X, xylem axis. **e,** DII-VENUS (blue) and *DR5::GFP* (yellow). **f–i,** DII-VENUS (**f, g**), mDII-VENUS (**h**) and *DR5::VENUS* (**i**) in the vegetative shoot apex (**f, h, i**) and hypocotyl (**g**). M, meristem; L, leaves; V, vasculature. Green or blue channel, VENUS; yellow channel in **e**, GFP; red channel, FM4-64 (**d, e**) or autofluorescence (**f–i**). In **f–i** the transmission channel has been added. Scale bars, 50 μ m.

in roots following exogenous auxin treatment (Fig. 3a, b and Supplementary Movie 1). Time-lapse confocal imaging revealed that the DII-VENUS signal was rapidly lost from all root tissues, whereas the signal in untreated roots remained stable (Fig. 3a and Supplementary Movie 1). Quantification of VENUS fluorescence in the root tip showed that a reduction in the DII-VENUS signal was detected minutes after auxin addition and the signal was abolished within 60 min (Fig. 3b). In contrast, an increase in the *DR5::VENUS* signal was first detected only after 120 min (Fig. 3b). This delay is due to post-transcriptional processes, because quantitative reverse transcription-PCR (qRT-PCR) detected *VENUS* messenger RNA minutes after auxin treatment (Fig. 3b). We could further show that the DII-VENUS degradation kinetics upon auxin treatment is very similar in different root tissues, but is slower in the root cap (Supplementary Fig. 8). These results indicate similar auxin sensitivity throughout the root except for a lower sensitivity in the root cap, as already suggested by the distribution of the TIR/AFBs (Supplementary Fig. 5a–d). We also observed that the global dynamics of degradation of DII-VENUS was similar in the vegetative and inflorescence SAM, with a minimal fluorescence reached after 1 h (Fig. 3c and Supplementary Fig. 9). We conclude that DII-VENUS responds almost immediately and similarly

to exogenous auxin application in various tissues. This observation strengthens our conclusion that DII-VENUS fluorescence is directly related to auxin levels in both shoot and root meristematic tissues but that co-receptor distribution needs to be considered.

Finally, we used DII-VENUS to follow changes in auxin distribution during developmental processes. Roots have been proposed to bend in response to gravity by accumulating auxin on the lower side of root apical tissues^{21,24,25}. Consistent with this model, induction of the *DR5* reporter occurs after 1.5–2 h in the lateral root cap (LRC) and epidermis on the lower side of the root (Supplementary Fig. 10)²¹. By contrast, within 30 min of a 90° gravity stimulus the DII-VENUS signal was entirely lost in these tissues on the lower side, whereas fluorescence was stable on the upper side (Fig. 4a). A decrease in DII-VENUS fluorescence was also observed in the cortex and endodermis on both sides of the root and to a lesser extent in vascular tissues. Hence, DII-VENUS

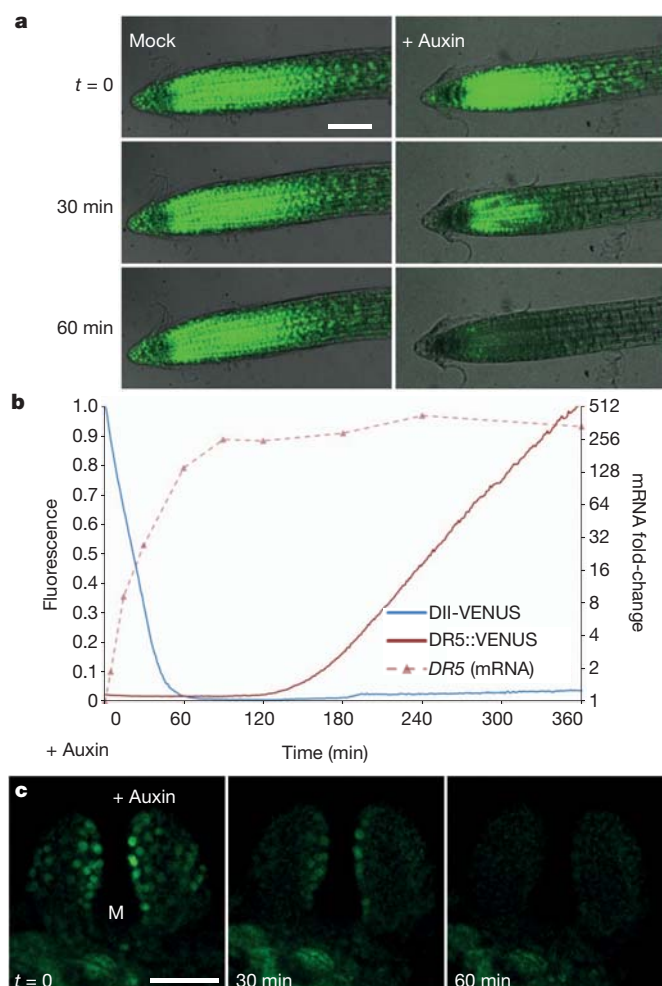


Figure 3 | DII-VENUS monitors changes in auxin response and distribution at high temporal resolution. **a**, Time-course of DII-VENUS fluorescence following either a mock or an auxin treatment (1 μ M 1-naphthaleneacetic acid; NAA). **b**, Quantification of DII-VENUS and *DR5::VENUS* fluorescence and of *DR5::VENUS* mRNA levels in root apices treated with 1 μ M NAA. **c**, Time-course of DII-VENUS fluorescence in the shoot apex upon 1 μ M indole-3-acetic acid treatment; images are projections of 10 confocal serial sections. Green channel, VENUS. In **a** the transmission channel has been added. Scale bars, 50 μ m.

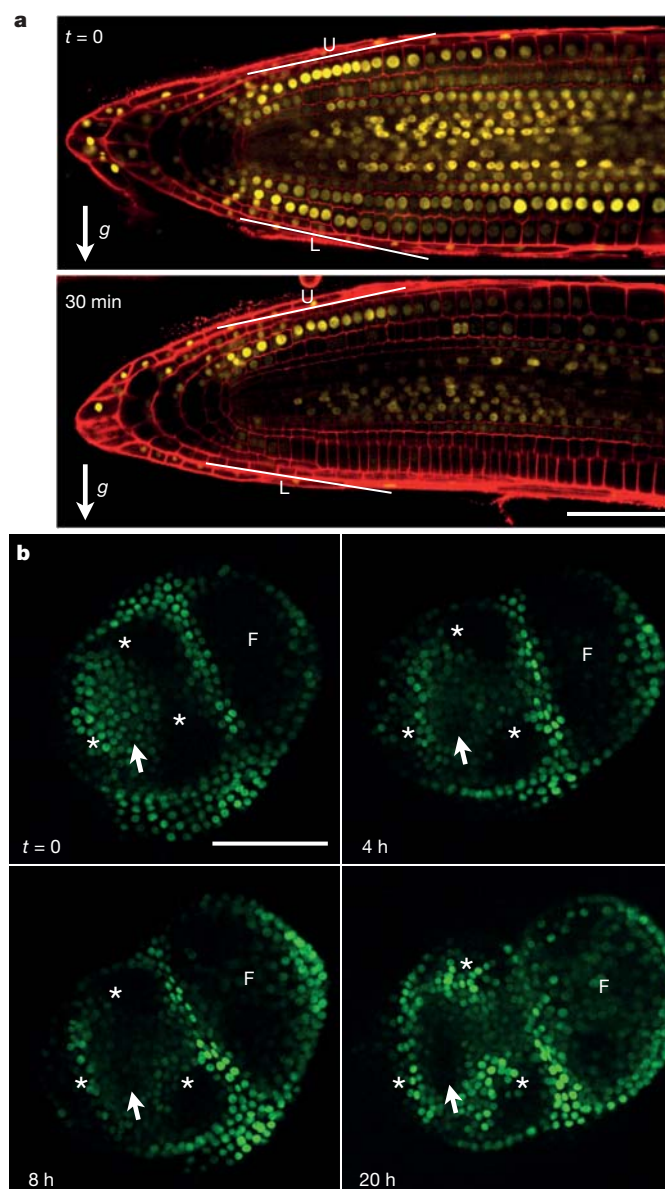


Figure 4 | DII-VENUS allows visualization of changes in auxin distribution during development. **a**, Changes in DII-VENUS fluorescence during a root gravitropic response; g, gravity vector; a line was drawn near equivalent cells in the LRC and epidermis on the lower (L) and upper (U) side. **b**, Auxin build-up during organ initiation at the shoot apex visualized using DII-VENUS; arrow, site of new organ initiation; asterisks, groups of nuclei showing notable changes in fluorescence; F, flower. Yellow or green channel, VENUS; red channel, propidium iodide. Scale bars, 50 μ m.

indicates that the changes in auxin concentration are not restricted to the LRC and epidermis during the gravitropic response. We next used DII-VENUS to follow auxin-dependent organ initiation at the SAM (Fig. 4b)^{26,27}. DII-VENUS allowed the visualization of the progressive build-up of auxin triggering the formation of a new organ (Fig. 4b). It also demonstrated significant redistribution of DII-VENUS fluorescence throughout the SAM, most probably reflecting the dynamics of auxin transport in the tissue²⁷. Taken together, our observations demonstrate that, during both shoot and root development, DII-VENUS detects dynamic changes in endogenous auxin distribution and responses that are more complex than previously thought. By demonstrating that DII-VENUS and DR5 fluorescence patterns are only partly complementary, we also provide evidence that the auxin signalling pathway has a key role in the spatial control of transcription in response to auxin during developmental processes. Finally, several other plant hormones have been shown to signal through degradation of key signalling regulators^{28–30}. Our work provides the foundation for building synthetic signalling sensors for different hormones to explore their role during development.

METHODS SUMMARY

Generation and characterization of DII-VENUS reporter plants. Aux/IAA degron sequences¹⁶ were fused to VENUS-NLS under the control of the 35S promoter (Supplementary Fig. 1) and recombined into gateway binary vector pH7m34GW (<http://gateway.psb.ugent.be/>). Plasmids were transformed into Col-0 plants by floral dipping.

Microscopy, live imaging and chemical treatments. All images were obtained using laser-scanning confocal microscope. Tissue organization was visualized using FM4-64, propidium iodide, transmission or chlorophyll autofluorescence. For live imaging of shoot meristems, plants were grown on the auxin transport inhibitor 1-N-naphthylphthalamic acid (NPA) to produce naked apices before transferring to a new medium without NPA. To generate maps of relative auxin distribution in the root meristem, a cellular grid was generated from a median optical section and for each cells fluorescence was extracted from the optical section cutting the nucleus at its most median part.

For analysis of chemically treated roots, 5-day-old DII-VENUS seedlings were transferred to media containing the chemicals at the stated concentration. For root live imaging, immediately following transfer, the seedlings were scanned every 2 to 5 min for 2 h. The fluorescence intensity over identical scanned portion of the root (corresponding approximately to the first 200 μm from the root tip) was extracted at each time point. For tissue-specific kinetics, fluorescence was extracted from groups of nuclei with nearly identical fluorescence in the different tissues. For the vegetative SAM, 5-day-old DII-VENUS seedlings were mounted in water after removing one cotyledon to allow observation of the vegetative shoot apex before treating with the stated concentration of auxin.

Pull-down assays. Pull-down assays were performed using a 35S:Flag-TIR1 line and biotinylated IAA28 peptide. The immunodetection of TIR1/AFB-Flag was performed with a 1:5,000 dilution of anti-Flag 2-peroxidase (HRP) antibody followed by chemiluminescent detection with ECL plus reagents.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 July; accepted 19 December 2011.

Published online 15 January 2012.

1. Friml, J. Auxin transport — shaping the plant. *Curr. Opin. Plant Biol.* **6**, 7–12 (2003).
2. Chapman, E. J. & Estelle, M. Mechanism of auxin-regulated gene expression in plants. *Annu. Rev. Genet.* **43**, 265–285 (2009).
3. Dharmasiri, N., Dharmasiri, S. & Estelle, M. The F-box protein TIR1 is an auxin receptor. *Nature* **435**, 441–445 (2005).
4. Dharmasiri, N. *et al.* Plant development is regulated by a family of auxin receptor F box proteins. *Dev. Cell* **9**, 109–119 (2005).
5. Tan, X. *et al.* Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**, 640–645 (2007).
6. Kepinski, S. & Leyser, O. The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* **435**, 446–451 (2005).
7. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nature Methods* **2**, 905–909 (2005).
8. Greenham, K. *et al.* The AFB4 auxin receptor is a negative regulator of auxin signaling in seedlings. *Curr. Biol.* **21**, 520–525 (2011).

9. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. ARF1, a transcription factor that binds to auxin response elements. *Science* **276**, 1865–1868 (1997).
10. Guilfoyle, T. J. & Hagen, G. Auxin response factors. *Curr. Opin. Plant Biol.* **10**, 453–460 (2007).
11. Sabatini, S. *et al.* An auxin-dependent distal organizer of pattern and polarity in the *Arabidopsis* root. *Cell* **99**, 463–472 (1999).
12. Ulmasov, T., Murfett, J., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**, 1963–1971 (1997).
13. Gray, W. M., Kepinski, S., Rouse, D., Leyser, O. & Estelle, M. Auxin regulates SCF^{TIR1}-dependent degradation of AUX/IAA proteins. *Nature* **414**, 271–276 (2001).
14. Ramos, J. A., Zenser, N., Leyser, O. & Callis, J. Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent. *Plant Cell* **13**, 2349–2360 (2001).
15. Zenser, N., Ellsmore, A., Leasure, C. & Callis, J. Auxin modulates the degradation rate of Aux/IAA proteins. *Proc. Natl Acad. Sci. USA* **98**, 11795–11800 (2001).
16. Dreher, K. A., Brown, J., Saw, R. E. & Callis, J. The *Arabidopsis* Aux/IAA protein family has diversified in degradation and auxin responsiveness. *Plant Cell* **18**, 699–714 (2006).
17. Abel, S., Oeller, P. W. & Theologis, A. Early auxin-induced genes encode short-lived nuclear proteins. *Proc. Natl Acad. Sci. USA* **91**, 326–330 (1994).
18. Vernoux, T. *et al.* The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* **7**, 508 (2011).
19. Parry, G. *et al.* Complex regulation of the TIR1/AFB family of auxin receptors. *Proc. Natl Acad. Sci. USA* **106**, 22540–22545 (2009).
20. Benková, E. *et al.* Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* **115**, 591–602 (2003).
21. Ottenschläger, I. *et al.* Gravity-regulated differential auxin transport from columella to lateral root cap cells. *Proc. Natl Acad. Sci. USA* **100**, 2987–2991 (2003).
22. Santuari, L. *et al.* Positional information by differential endocytosis splits auxin response to drive *Arabidopsis* root meristem growth. *Curr. Biol.* **21**, 1918–1923 (2011).
23. Petersson, S. V. *et al.* An auxin gradient and maximum in the *Arabidopsis* root apex shown by high-resolution cell-specific analysis of IAA distribution and synthesis. *Plant Cell* **21**, 1659–1668 (2009).
24. Boonsirichai, K., Guan, C., Chen, R. & Masson, P. H. Root gravitropism: an experimental tool to investigate basic cellular and molecular processes underlying mechanosensing and signal transmission in plants. *Annu. Rev. Plant Biol.* **53**, 421–447 (2002).
25. Boonsirichai, K., Sedbrook, J. C., Chen, R., Gilroy, S. & Masson, P. H. ALTERED RESPONSE TO GRAVITY is a peripheral membrane protein that modulates gravity-induced cytoplasmic alkalization and lateral auxin transport in plant statocytes. *Plant Cell* **15**, 2612–2625 (2003).
26. Reinhardt, D. *et al.* Regulation of phyllotaxis by polar auxin transport. *Nature* **426**, 255–260 (2003).
27. Heisler, M. G. *et al.* Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. *Curr. Biol.* **15**, 1899–1911 (2005).
28. Silverstone, A. L. *et al.* Repressing a repressor: gibberellin-induced rapid reduction of the RGA protein in *Arabidopsis*. *Plant Cell* **13**, 1555–1566 (2001).
29. Fu, X. & Harberd, N. P. Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature* **421**, 740–743 (2003).
30. Santner, A. & Estelle, M. Recent advances and emerging trends in plant hormone signalling. *Nature* **459**, 1071–1078 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Erktan and C. Cellier for help with marker expression analysis; J. Neve, A. Miyawaki, M. Heisler and M. Estelle for providing the 35S::Flag-TIR1 line, VENUS complementary DNA, DR5::VENUS plasmids and TIR/AFB GUS lines, respectively; the PLATIM for access to confocal microscopes; F. Parcy, O. Hamant, A. Boudaoud and P. Das for discussions. T.V. was supported by the Human Frontier Science Program Organization (CDA 0047/2007 HFSP) and the Agence National de la Recherche (ANR-07-JCJC-0115 and EraSysBio+ iSAM). D.M.W., A.L. and M.J.B. acknowledge the support of the Biotechnology and Biological Sciences Research Council (BBSRC) and Engineering and Physical Sciences Research Council (EPSRC) funding to the Centre for Plant Integrative Biology (CPiB); BBSRC grants BB/F013981/1 and BB/F007418/1 to S.K.; BBSRC Professorial Research Fellowship funding to D.M.W. & M.J.B.; and Belgian Scientific policy (BELSPO contract BARN) to A.L., T.B. and M.J.B.

Author Contributions T.V. designed the DII-VENUS tool. G.B., M.O. and T.V. engineered and characterized DII-VENUS transgenic lines. D.M.W., G.B., A.L. and V.M. quantified the spatial and temporal dynamics of DII-VENUS. A.H.B. did the pull-down assay. T.V. and M.J.B. designed the experiments with the help of T.B., S.K. and J.T. T.V. and M.J.B. analysed the data and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Seed for the lines described in this study have been deposited at the Nottingham *Arabidopsis* Stock Centre. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.V. (teva.vernoux@ens-lyon.fr).

METHODS

Plant material, growth conditions and plant treatments. All transgenic plants were generated in the Columbia ecotype (Col-0). The *tir1 afb1 afb2 afb3* quadruple mutant, DR5::GFP line and TIR1/AFB1–AFB3 GUS translational fusions have been described^{4,19,31}. DR5::VENUS transgenic plants were generated by transforming a DR5::VENUS plasmid²⁷ by floral dipping³². Plants were cultivated *in vitro* on MS medium supplemented with 1% sucrose at 22 °C and under long-day conditions (16 h light/8 h darkness). For analysis on roots, the chemical treatments were done on 5-day-old plants by transferring them to liquid MS supplemented with the chemicals or on an MS agar supplemented with the chemicals for root live imaging. Indole-3-acetic acid (Sigma) or 1-naphthaleneacetic acid (NAA; Sigma) was dissolved in ethanol and used at the indicated concentration. MG132 and clasto-lactacystin- β -lactone (lactacystin; Sigma) were dissolved in dimethylsulphoxide (DMSO) and used at the final concentration of 50 μ M for 2.5 h or 20 μ M for 8 h respectively. For MG132/indole-3-acetic acid co-treatments, plants were pretreated with MG132 for 1.5 h before adding indole-3-acetic acid. For analysis on the vegetative shoot apex, 5-day-old seedlings were used after removing one cotyledon to allow observation of the vegetative shoot apex. Seedlings were mounted in water and treatments were done by replacing by capillarity the water with indole-3-acetic acid at the indicated concentration. For the inflorescence apex, the plants were transferred to water containing indole-3-acetic acid at the indicated concentration.

Generation of DII-VENUS transgenic plants. The DII-VENUS binary vectors were generated using Gateway technology and following the Multisite Gateway three-fragment vector construction kit protocol (Invitrogen). To generate the different versions of the DII-VENUS sensor (Supplementary Fig. 1), we used the region of IAA8, IAA9 and IAA28 starting from the conserved lysine up to the end of domain II (IAA8, amino acid positions 107–178; IAA9, 120–195; IAA28, 28–61; Supplementary Fig. 1). IAA8, IAA9 and IAA28 were chosen because their basal half-lives were potentially long enough¹⁶ (ranging from 15–20 min for IAA8 and IAA9 to 80 min for IAA28) to allow for the maturation of the fast-maturing YFP variant VENUS and thus for visualization of its fluorescence. We cloned the IAA8, IAA9 and IAA28 cDNAs by standard RT–PCR from inflorescence mRNA. The different wild-type sequences were then amplified by PCR (see Supplementary Table 1 for primers) and cloned in pDONR 221 by recombination. We then mutated the conserved lysine (K to R mutation) by introducing this mutation in the forward primers (sequence in bold replaced by AGA: see Supplementary Table 1). To generate mDII-VENUS, site-directed mutagenesis (using standard inverted PCR procedures) was used to introduce the P53L mutation in the wild-type IAA28 sequence in pDONR 221 (Supplementary Fig. 1)³³. The sequence of VENUS fused to the N7 nuclear localization signal³⁴ was amplified by PCR (see Supplementary Table 1 for primers) from a VENUS-N7 sequence cloned in pBG36 and cloned into pDONR P2R-P3 by recombination. Finally Aux/IAA-derived sequences were fused in-frame to VENUS-N7 (Supplementary Fig. 1) and put under the control of the strong constitutive CaMV 35S promoter, using a 35S promoter cloned in pDONR P4-P1R and recombination into the gateway-compatible pH7m34GW binary vector³⁵ (hygromycin resistance). The different plasmids were then introduced in plants by floral dipping³².

Confocal microscopy, live imaging and quantification of fluorescence. Imaging was performed either on either a LSM-510 laser-scanning confocal microscope (Zeiss), a SP5 spectral detection confocal microscope (Leica) or an Eclipse Ti 2000 laser-scanning confocal microscope (Nikon). For visualization of the root organization the roots were stained either with FM4-64 (Invitrogen) as previously described³⁶ or propidium iodide (Sigma). To quantify fluorescence with cellular resolution in DII-VENUS and mDII-VENUS root meristems and generate maps of relative auxin using DII-VENUS, serial optical sections were obtained. A cellular grid was generated from the propidium iodide channel of the most median optical section using MerrySim³⁷. For each cell defined in the grid, we then selected the optical section passing through the centre of each nuclei. Fluorescence was then summed inside the corresponding cell from that section. The loss of fluorescence due to tissue absorbance was also estimated using the spatial distribution of the propidium iodide channel and used to correct the fluorescence values. To obtain the changes in auxin levels along the root axis in the different tissues, fluorescence distribution was extracted and curvatures were smoothed using a Gaussian kernel with a sigma value of 5.

For root live imaging, immediately after the beginning of the treatment, the seedlings were scanned every 2 min for 2 h to follow the evolution of the DII-VENUS signal. To quantify fluorescence in the root tip, the average fluorescence intensity over identical scanned portion of the root (corresponding approximately to the first 200 μ m from the root tip) was extracted using EZ-C1 software (v3.9, Nikon) and the values analysed using Microsoft Excel. For tissue specific kinetics, the seedlings were scanned every 5 min and fluorescence was extracted from groups of nuclei with nearly identical fluorescence (variations < 30%) in the different tissues. The fluorescence intensity of nuclei was extracted from the different

tissues using the ROI tool of Fiji software (<http://fiji.sc/wiki/index.php/Fiji>) and the values analysed using Microsoft Excel. For dose-dependent quantification of DII-VENUS signal upon auxin treatment, fluorescence was measured 1 h after treatment using two and three roots for wild-type and *tir1 afb1 afb2 afb3*, respectively.

For the vegetative shoot apex, seedlings were mounted into water in between slide and cover slip. For live imaging of vegetative shoot apex, five plants were treated after observation at $t = 0$ and followed over 90 min. For the inflorescence apex, observation was performed as described³⁸. Live imaging of the inflorescence apex was performed on plants grown on the auxin transport inhibitor NPA then transferred to a new medium without NPA as previously described³⁹. Initiation of a new organ was confirmed a posteriori by visual inspection of the apex.

Root growth analysis, gravitropic assays and flower production rate. For analysis of root growth, plants were grown near-vertically and root length was measured at the indicated time. For gravitropic assays, plants were grown as detailed previously⁴⁰ and imaged at 30 min intervals following a 90° gravi-stimulus. Root tip angle was measured using modified RootTrace software⁴¹ (<http://www.cpi.ac.uk/tools-resources/roottrace>). Estimation of flower production rate was done as described⁴².

IAA28-Flag-TIR1 pull-down assays. To generate the 35S::Flag-TIR1 transgenic *Arabidopsis* line a plant expression vector containing a 3×Flag was first created by annealing complementary 101-base-pairs oligonucleotides including the 3×Flag coding sequences (see Supplementary Table 1) and cloning this fragment into XbaI and SalI sites of the vector pFP101. The Gateway C1 cassette (Invitrogen Gateway vector conversion reagent system) was then introduced into this plasmid by blunt-end ligation following SalI digestion and end-filling by Klenow reaction to create the destination vector pFP3FLAGSI. The full-length coding sequence for TIR1 was amplified from an *Arabidopsis* cDNA library using Gateway-compatible primers (see Supplementary Table 1) and incorporated into the Gateway donor vector pDONR207 by BP reaction (Invitrogen). The TIR1 coding sequence was then incorporated into pFP3FLAGSI via a Gateway vector. LR reaction was used to form plasmid pFP3FLAGSI-TIR1. Wild-type *Arabidopsis* plants were subsequently transformed using the floral dip method³² and homozygous lines with single-site were selected from the T₃ generation. Extracts of 10-day-old 35S::Flag-TIR1 seedlings were made as described previously⁶ and used in pull-down assays by combining 2.5 mg of crude extract with 5 μ g of biotinylated IAA28 domain II peptide (biotinyl-NH-EVAPVVGWPPVRSRRN-COOH, synthesized by Thermo Scientific), and 65 μ l 50% streptavidin-agarose suspension. The assays were incubated for 1 h at 4 °C with mixing then washed three times for 5 min in extraction buffer (0.15 M NaCl, 0.5% Nonidet P40, 0.1 M Tris-HCl pH 7.5, containing 1 mM phenylmethylsulphonyl fluoride, 1 μ M dithiothreitol, 10 μ M MG132) containing the appropriate auxin treatment. The final processing of the pull-down assays including electrophoresis and western transfer were performed as described previously⁶. The immunodetection of TIR1/AFB-Flag was performed with a 1:5,000 dilution of anti-Flag 2-peroxidase (HRP) antibody (Sigma) followed by chemiluminescent detection with ECL plus reagents (Amersham).

Transcript profiling. Total RNA was extracted from roots using an RNeasy Plant Micro Kit (Qiagen), including on-column DNase digestion to eliminate genomic DNA from the samples. A 500-ng aliquot of RNA was reverse-transcribed using a Transcriptor First Strand cDNA synthesis kit (Roche) and anchored-oligo (dT)18. Real-time qPCRs were performed on a Roche Light Cycler 480 system using the prevalidated single hydrolysis probes, Sensimix probe master mix (Quantace), and gene-specific primers (see Supplementary Table 1).

1. Friml, J. *et al.* Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* **426**, 147–153 (2003).
2. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
3. Rogg, L. E., Lasswell, J. & Bartel, B. A gain-of-function mutation in IAA28 suppresses lateral root development. *Plant Cell* **13**, 465–480 (2001).
4. Cutler, S. R., Ehrhardt, D. W., Griffiths, J. S. & Somerville, C. R. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of *Arabidopsis* at a high frequency. *Proc. Natl Acad. Sci. USA* **97**, 3718–3723 (2000).
5. Karimi, M., De Meyer, B. & Hilson, P. Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103–105 (2005).
6. Levesque, M. P. *et al.* Whole-genome analysis of the SHORT-ROOT developmental pathway in *Arabidopsis*. *PLoS Biol.* **4**, e143 (2006).
7. Barbier de Reuille, P., Bohn-Courseau, I., Godin, C. & Traas, J. A protocol to analyse cellular dynamics during plant development. *Plant J.* **44**, 1045–1053 (2005).
8. Fernandez, R. *et al.* Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. *Nature Methods* **7**, 547–553 (2010).
9. Grandjean, O. *et al.* *In vivo* analysis of cell division, cell growth, and differentiation at the shoot apical meristem in *Arabidopsis*. *Plant Cell* **16**, 74–87 (2004).
10. Holman, T. J. *et al.* Statistical evaluation of transcriptomic data generated using the Affymetrix one-cycle, two-cycle and IVT-Express RNA labelling protocols with the *Arabidopsis* ATH1 microarray. *Plant Methods* **6**, 9 (2010).

41. French, A., Ubeda-Tomas, S., Holman, T. J., Bennett, M. J. & Pridmore, T. High-throughput quantification of root growth using a novel image-analysis tool. *Plant Physiol.* **150**, 1784–1795 (2009).
42. Lohmann, D. *et al.* SLOW MOTION is required for within-plant auxin homeostasis and normal timing of lateral organ initiation at the shoot meristem in *Arabidopsis*. *Plant Cell* **22**, 335–348 (2010).

Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins

Philipp Engel^{1†*}, Arnaud Goepfert^{1,2*}, Frédéric V. Stanger^{1,2}, Alexander Harms¹, Alexander Schmidt³, Tilman Schirmer² & Christoph Dehio¹

Fic proteins that are defined by the ubiquitous FIC (filamentation induced by cyclic AMP) domain are known to catalyse adenylylation (also called AMPylation); that is, the transfer of AMP onto a target protein. In mammalian cells, adenylylation of small GTPases through Fic proteins injected by pathogenic bacteria can cause collapse of the actin cytoskeleton and cell death^{1,2}. It is unknown how this potentially deleterious adenylylation activity is regulated in the widespread Fic proteins that are found in all domains of life and that are thought to have critical roles in intrinsic signalling processes. Here we show that FIC-domain-mediated adenylylation is controlled by a conserved mechanism of ATP-binding-site obstruction that involves an inhibitory α -helix (α_{inh}) with a conserved (S/T)XXE(G/N) motif, and that in this mechanism the invariable glutamate competes with ATP γ -phosphate binding. Consistent with this, FIC-domain-mediated growth arrest of bacteria by the VbhT toxin of *Bartonella schoenbuchensis* is intermolecularly repressed by the VbhA antitoxin through tight binding of its α_{inh} to the FIC domain of VbhT, as shown by structure and function analysis. Furthermore, structural comparisons with other bacterial Fic proteins, such as Fic of *Neisseria meningitidis* and of *Shewanella oneidensis*, show that α_{inh} frequently constitutes an amino-terminal or carboxy-terminal extension to the FIC domain, respectively, partially obstructing the ATP binding site in an intramolecular manner. After mutation of the inhibitory motif in various Fic proteins, including the human homologue FICD (also known as HYPE), adenylylation activity is considerably boosted, consistent with the anticipated relief of inhibition. Structural homology modelling of all annotated Fic proteins indicates that inhibition by α_{inh} is universal and conserved through evolution, as the inhibitory motif is present in ~90% of all putatively adenylylation-active FIC domains, including examples from all domains of life and from viruses. Future studies should reveal how intrinsic or extrinsic factors modulate adenylylation activity by weakening the interaction of α_{inh} with the FIC active site.

In two Fic proteins, IbpA and VopS, that are translocated by pathogenic bacteria into host cells, the ubiquitous FIC domain has been shown to catalyse adenylylation^{1–4}. The crystal structure of the effector domain IbpA(FIC2) in complex with its adenylylated host target Cdc42 has been reported⁵ and a catalytic mechanism has been proposed^{5,6}. IbpA- or VopS-mediated adenylylation of Rho-family GTPases abolishes downstream signalling in human cells and, thus, causes actin cytoskeleton collapse and cell death^{1,2}. By contrast, overexpression of a human Fic protein with similar target specificity, HYPE, had only a marginal effect¹. This suggests that the potentially deleterious adenylylation activity is tightly regulated in HYPE and probably in most of the almost 3,000 Fic proteins that are proposed to have important roles in intrinsic signalling processes in bacteria, archaea and eukaryotes.

VbhT is a bacterial Fic protein of the mammalian pathogen *B. schoenbuchensis*^{7,8}. It is composed of an N-terminal FIC domain

and a C-terminal BID domain (Fig. 1a). The BID domain facilitates protein translocation into mammalian or bacterial target cells through a type IV secretion system or conjugation machinery⁹, respectively, but the target cell and functional role of VbhT are unknown. VbhT arrests growth when expressed in *Escherichia coli* (Fig. 1c and Supplementary Fig. 1a). Growth arrest is repressed by mutation of histidine to alanine (VbhT(H136A)) in the conserved FIC motif HXFX(D/E)GNRXXR. In other Fic proteins, this signature motif has been shown to be essential for target protein adenylylation activity^{1,2,10}, therefore suggesting that toxicity is related to adenylylation of endogenous proteins. Indeed, wild-type VbhT, but not VbhT(H136A), catalysed *in vitro* adenylylation of a putative *E. coli* target protein of approximately 80 kilodaltons (kDa) (Fig. 1d). Furthermore, *E. coli* cells showed filamentation after expression of wild-type VbhT, but not of VbhT(H136A) (Supplementary Fig. 2). A similar phenotype has been described for what is thought to be a hyperactive mutant of the *E. coli* Fic protein¹¹. Co-expression of VbhT with VbhA, encoded by the small open reading frame *vbhA*

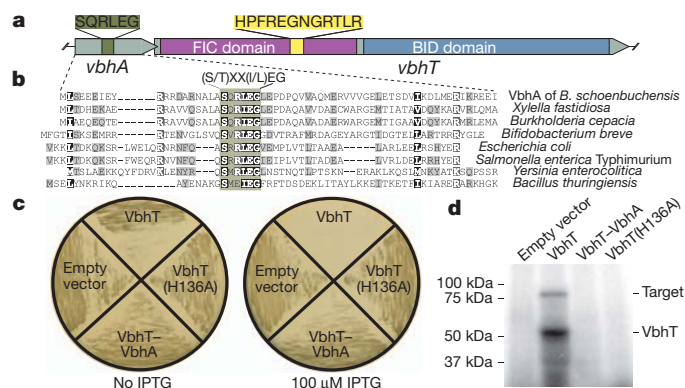


Figure 1 | The small protein VbhA represses the toxic effect (growth arrest) that is mediated by the adenylylation activity of VbhT in *E. coli*. **a**, Genetic organization of the type-IV-secretion-system-associated locus of *B. schoenbuchensis*, which is composed of the overlapping *vbhA* and *vbhT* genes. The FIC and BID domains that are encoded by *vbhT* are shown in different colours. Protein translations of the regions encoding the conserved motif of VbhA and the FIC motif of VbhT are depicted. **b**, Protein alignment of VbhA and a subset of the 158 homologues that are encoded upstream of *fic* loci in different bacteria (see also Supplementary Fig. 4). The most conserved region shows a (S/T)XX(I/L)EG consensus. Sequence accessions and strain designations are given in Supplementary Fig. 4. **c**, Growth of *E. coli* on lysogeny broth (LB) plates after IPTG-induced expression of different VbhT and VbhA constructs. Expression of VbhT shows a toxic effect, whereas bacterial growth is not affected when VbhT(H136A) is expressed or when VbhT and VbhA are co-expressed (VbhT–VbhA). All *E. coli* strains revealed normal growth without induction. Growth curve experiments in LB gave analogous results (Supplementary Fig. 1). **d**, Adenylylation assay with crude cell lysates of *E. coli* ectopically expressing the same constructs as in panel c.

¹Focal Area Infection Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. ²Core Program Structural Biology and Biophysics, Biozentrum, University of Basel, CH-4056 Basel, Switzerland.

³Proteomics Core Facility, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [†]Present address: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520-8106, USA.

*These authors contributed equally to this work.

immediately upstream of *vbhT* (Fig. 1a), completely repressed VbhT toxicity, as shown by wild-type-like bacterial growth, normal cell morphology, and inhibition of VbhT-dependent adenylylation (Fig. 1 and Supplementary Figs 1a and 2). We also observed VbhT-mediated toxicity and its repression by VbhA in *B. schoenbuchensis*, the natural carrier of this toxin and antitoxin, and in the related species *Bartonella henselae* (Supplementary Fig. 3).

The inhibitory action of VbhA on the VbhT toxin, and the genetic organization of the respective genes in an operon are reminiscent of toxin–antitoxin modules that are found in many bacterial genomes, often associated with mobile genetic elements¹². A comprehensive analysis of the upstream region of FIC-domain-encoding genes (PFAM pf02661) identified 158 bacterial *vbhA* homologues that probably function as antitoxins. Although the sequences are rather diverse, a central (S/T)XX(I/L)EG motif is conspicuous (Fig. 1b and Supplementary Fig. 4). The high-resolution (1.5 Å) crystal structure of VbhA in complex with the FIC domain of VbhT (VbhT(FIC)) (Supplementary Table 1) shows that VbhA is folded into three anti-parallel helices that tightly embrace VbhT(FIC) (Fig. 2a) with the N-terminal helix (α_{inh}), adopting a location that is analogous but distinct to that of the antitoxin Phd in its complex with Doc¹³ (Supplementary Fig. 5). Doc is a Fic protein with a degenerate, probably adenylation-incompetent FIC motif⁴ that may have adopted another toxic activity (Supplementary Information, section 1). The VbhA antitoxin motif locates to the C-terminal part of α_{inh} and is positioned close to the putative ATP-binding site¹⁰ at the N-cap of the helix that follows the active loop of VbhT(FIC). This suggests that the antitoxin competes with ATP binding. VbhA residues Ser 20 and Glu 24 of the inhibitory motif form a hydrogen bond and a salt bridge, respectively, with the conserved Arg147 of VbhT following the active loop.

Intriguingly, structural comparison with other bacterial Fic proteins of known fold (Fic proteins from *S. oneidensis* (SoFic)^{14,15} and from *N. meningitidis* (NmFic)¹⁴ (Fig. 2b, c), and from *Bacteroides thetaiotaomicron* (BtFic)¹⁴ and from *Helicobacter pylori* (HpFic)¹⁴ (Supplementary Fig. 6)) reveals that a structural equivalent of α_{inh} can be part of the FIC domain fold itself. Moreover, these proteins also show the SXXXE(G/N) inhibitory motifs that are, with respect to the FIC active site, arranged exactly as in the VbhA–VbhT(FIC) complex. Along the polypeptide chain, however, these α_{inh} occur at two distinct locations either in the N-terminal part (SoFic and BtFic) or at the C terminus (NmFic and HpFic). Thus, Fic proteins containing α_{inh} can be grouped into three classes (Fig. 2d) depending on whether α_{inh} is provided by an interacting antitoxin (class I) or whether it is part of the FIC fold as an N-terminal helix (class II) or a C-terminal helix (class III).

To investigate the distribution of class II and class III Fic proteins, we predicted the structures of all PFAM FIC domain entries by homology modelling (Supplementary Information, section 1). Including the class I proteins that are identified above, two-thirds of all Fic proteins were classified (Fig. 2e and Supplementary Table 2), with a strong dominance of class II, the class to which human HYPE belongs (Supplementary Tables 3 and 4). The proportion of classified Fic proteins increases to 90% when considering only adenylation-competent Fic proteins that are defined by compliance with the HXFX(D/E)GNRXXR motif (Supplementary Text). This suggests co-evolution of catalytic and inhibitory function. The inhibition motifs that are derived from Psi-Blast (class I) and structural predictions (classes II and III) are shown in Fig. 2f with the overall consensus being (S/T)XXXE(G/N). The strict conservation of the glutamate is striking and indicates that the observed ionic interaction with the second arginine of the conserved FIC motif (Figs 2a–c) is crucial for inhibition. Phylogenetic distribution of class I

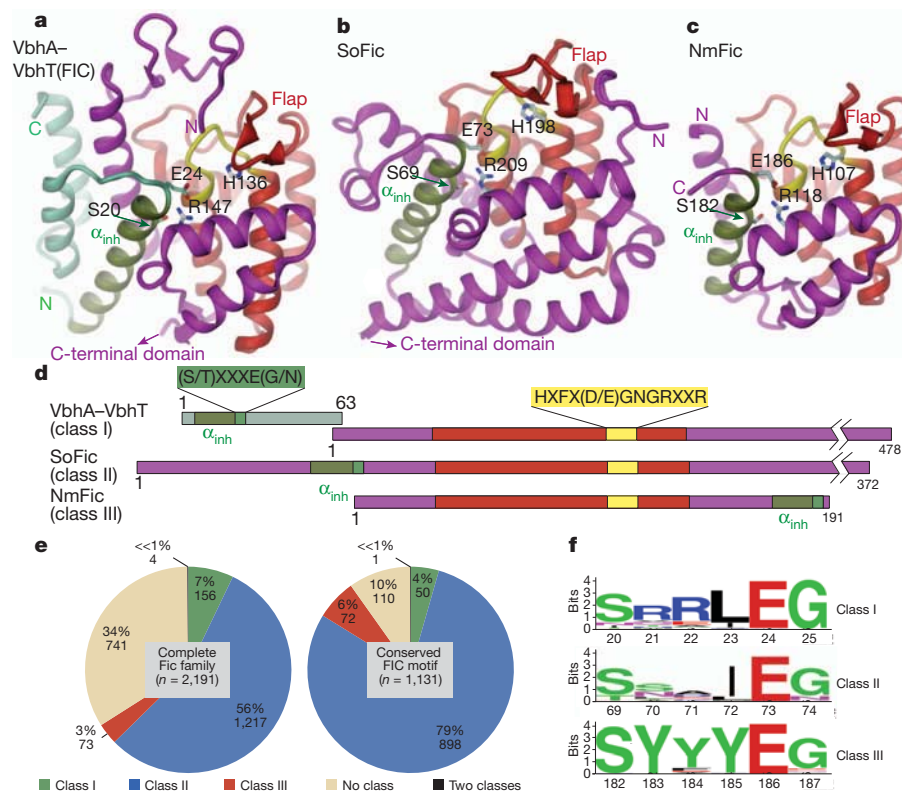


Figure 2 | Structures and classification of Fic proteins according to the position of the inhibitory motif along the polypeptide chain. Structures are shown in cartoon (FIC core as defined by PFAM, red; active site loop with histidine, yellow; inhibitory helix (α_{inh}) with SXXXE(G/N) motif, green) with important residues in full. **a**, Complex of VbhT(FIC) with antitoxin VbhA (green). **b**, SoFic (from *S. oneidensis*, PDB code 3EQX) with C-terminal domain omitted. **c**, NmFic (from *N. meningitidis*, PDB code 2G03). **d**, Linear

organization of motifs in the proteins presented in panels a–c. **e**, Distribution of the three Fic protein classes for the entire family in the PFAM database (left) and the subset of Fic proteins with the conserved FIC motif HXFX(D/E)GNRXXR that is likely to convey adenylation activity (right; see also Supplementary Information, section 1). **f**, Sequence profiles for the inhibition site of the three Fic protein classes.

and class III Fic proteins indicates that each is of monophyletic origin (Supplementary Fig. 7). Fic proteins with a degenerate FIC motif are dispersed over the tree, with the exception of the large cluster of Doc-like toxins. This suggests that there is recurrent degeneration of the conserved FIC motif with concomitant loss of adenylation activity. Consistent with this, deterioration of the FIC motif seems to correlate with the absence of a recognizable inhibitory motif (Supplementary Fig. 7). Fic proteins with a degenerate FIC motif may display catalytic activities different from adenylation, such as phosphocholination, as reported for the *Legionella pneumophila* effector AnkX¹⁶.

Owing to our discovery of the prevalence of the inhibitory motif in Fic proteins, we carried out a detailed analysis of its functional and structural role. For this, NmFic (class III, Fig. 2c) was chosen, as it is the smallest active Fic family protein with known crystallization condition (Supplementary Table 1). As reported before⁶, NmFic exhibits *in vitro* auto-adenylation activity (Fig. 3d). The acceptor site was traced to Y183 of α_{inh} by mass spectrometry (Supplementary Information, section 1, and Supplementary Fig. 8). On the basis of the location of Y183 relative to the active site, auto-adenylation is probably catalysed intermolecularly after partial unfolding or detachment of α_{inh} . Addition of *E. coli* lysate to NmFic did not reveal additional bands on the autoradiograph, and this indicated that there are no NmFic targets in *E. coli* or that the activity of NmFic is inhibited. The latter was shown to be correct as mutation of the inhibitory motif (S182A/E186A, NmFic(SE/AA)) resulted in transfer of radioactivity onto an ~80-kDa *E. coli* protein and enhanced auto-adenylation with an additional acceptor site (Y188; Fig. 3d and Supplementary Fig. 9). Deletion of the entire α_{inh} helix (NmFic(Δ 8)) led to similar target protein adenylation, proving that the activity resides in the FIC domain core. However, only weak auto-adenylation was apparent owing to the lack of the acceptor tyrosines in this deletion mutant (Fig. 3d and Supplementary Fig. 8c).

To investigate the inhibitory mechanism, crystals of NmFic proteins were soaked with the non-hydrolysable ATP analogue adenylyl

imidodiphosphate (AMPPNP) (Supplementary Table 1). The NmFic-AMPPNP structure revealed nucleotide binding but with the γ -phosphate disordered (Fig. 3a). Notably, the orientation of the α -phosphate seems to be non-productive, as the position that is in line with the scissile P α -O3 α bond is occluded by H107 and N113 (Supplementary Fig. 10a and Supplementary Movie). To reveal the situation in an inhibition-relieved mutant, the structure of NmFic(SE/AA) was determined to 3.0 Å (Supplementary Fig. 11). Electron density was lacking for α_{inh} , indicating disorder, whereas the nucleotide conformation was well defined. In the NmFic(Δ 8)-AMPPNP structure, the same nucleotide conformation was observed (Fig. 3b) and, owing to its high resolution (1.7 Å), the structural basis for the observed relief of inhibition in these mutants became evident. Whereas the adenosine moiety adopts the same position as in the wild-type, the γ -phosphate of the nucleotide is bound to R118, occupying the same position as the carboxylate of the inhibitory E186 in the wild type. As a consequence, the α -phosphate is found to be re-oriented, and the new orientation permits in line attack of a target side chain onto the α -phosphorus to accomplish AMP transfer (Fig. 3c, Supplementary Fig. 10 and Supplementary Movie).

The exact role of the inhibitory glutamate was investigated further by mutagenesis of Fic proteins from the three regulatory classes. In wild-type NmFic, C α and C β of the glutamate are close to the position that is attained by the γ -phosphate position in NmFic(Δ 8) (Fig. 3c). Still, an E to G single point mutant may provide sufficient space and main-chain flexibility to allow γ -phosphate binding. Indeed, similar to NmFic(Δ 8) and NmFic(SE/AA), the mutant NmFic(E186G) resulted in growth inhibition of *E. coli* (Supplementary Fig. 1b). Likewise, co-expression of VbhT(FIC) and VbhA(E24G), as representatives for class I, caused *E. coli* growth defects (Supplementary Fig. 1a). We also included SoFic, a bacterial class II protein, in this analysis. Consistent with the effects of E to G single point mutants in NmFic and VbhA, mutant SoFic(E73G) revealed a negative effect on *E. coli* growth

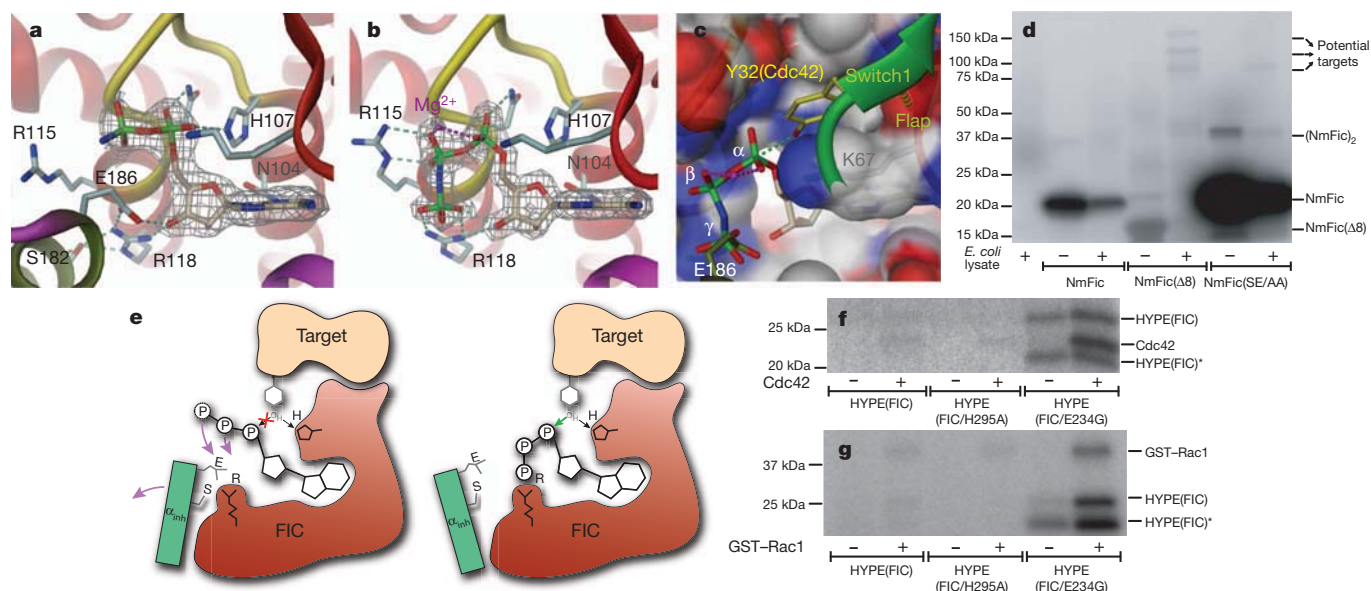


Figure 3 | Structure and function of wild-type and mutant NmFic reveals a general inhibition mechanism corroborated by HYPE protein analysis.

a, Active site of NmFic with bound ATP analogue AMPPNP. The γ -phosphate seems to be disordered and has not been modelled. **b**, Active site of NmFic(Δ 8) with bound AMPPNP and Mg²⁺. The γ -phosphate occupies the position taken by E186 of the SXXXEG motif in the wild-type (as shown in **a**). Also shown are 2F_o - F_c maps that are contoured at 1.2 σ and cover the ligands. **c**, Surface representation of the active site of NmFic(Δ 8) with modifiable Y32 (yellow) of the Cdc42 switch1 loop (green) in a position that is obtained from a superposition of the active site loops of NmFic over the IbpA(FIC2)-Cdc42 complex⁵. **d**, Autoradiography of an SDS gel after incubation of wild-type and mutant NmFic with [α -³²P]ATP in the presence or absence of *E. coli* lysate. The

mutants, but not the wild-type, catalyse AMP transfer onto *E. coli* target proteins. **e**, Scheme of a general inhibition mechanism for Fic proteins. The α_{inh} helix (green) with the (S/T)XXXE(G/N) motif prevents productive ATP binding. It is only after dissociation of the helix that the ATP γ -phosphate attains the position close to a conserved arginine (indicated by “R”) of the FIC motif. This is accompanied by reorientation of the α -phosphate to allow in-line attack of the target hydroxyl side chain after proton transfer to the active histidine as proposed before⁵. **f**, **g**, Autoradiography after incubation of HYPE(FIC) with α -³²P-ATP in the presence or absence of Cdc42 (**f**) or GST-tagged Rac1 (GST-Rac1) (**g**). HYPE(FIC/E234G) shows enhanced auto-adenylation and target adenylation. HYPE(FIC)* denotes a degradation product of HYPE(FIC).

(Supplementary Fig. 1c) and strongly enhanced auto-adenylation (Supplementary Fig. 12). A similar marked effect on auto-adenylation was observed after mutation of the corresponding residue in human HYPE (HYPE(FIC/E234G)), another class II protein (Fig. 3f, g). It is known that HYPE catalyses Cdc42 and Rac1 adenylation^{1,4}. The rather low adenylation activity of wild-type HYPE on these substrates is markedly enhanced in the HYPE(FIC/E234G) mutant (Fig. 3f, g), demonstrating that the relief of inhibition by α_{inh} enhances not only auto-adenylation but also AMP transfer onto bona fide protein targets.

From this study a general mechanism for the inhibition of the FIC-domain-mediated adenylation has emerged that invokes a glutamate finger from α_{inh} to inter- or intramolecularly block part of the ATP binding site (Fig. 3e). Dissociation of the toxin-antitoxin complex (class I) or intramolecular loosening of the contact between α_{inh} and the FIC domain active site (classes II and III) allows ordered binding of the entire ATP moiety with the α -phosphate in an orientation that is productive for accepting an incoming target hydroxyl side chain.

Class I proteins might exert functions similar to the classical bacterial toxin-antitoxin modules¹², whereas class II and III Fic proteins seem to regulate intrinsic cellular functions that are related to physiological adaptation and cell homeostasis. Conservation of class II Fic proteins in all three kingdoms of life (Supplementary Table 2) emphasizes the important role of the regulatory mechanisms described here. How the adenylation activity is activated by weakening the interaction of α_{inh} with the FIC active site in class II and class III Fic proteins is one of many important questions for future research. We also anticipate our study to be a starting point for rational approaches to modulate the adenylation activities of Fic proteins—approaches that should aid in elucidating the diverse biological functions of these widespread signalling proteins.

METHODS SUMMARY

Toxicity experiments were performed with *E. coli* strain MG1655, which encodes an IPTG (isopropyl- β -D-thiogalactoside)-inducible T7 polymerase (strain AB472). Protein expression was controlled by the addition of either 100 μ M IPTG (induction) or 1% glucose (repression). A PSI-BLAST search for homologues with VbhA of *B. schoenbuchensis* was conducted to identify Fic proteins belonging to class I. Class II and class III proteins were classified by structure predictions using the program HHpred¹⁷. We analysed all 2,189 proteins of the FIC PFAM family (pf02661, release 24). Adenylation assays were carried out using bacterial crude cell lysates for VbhT constructs or purified proteins for NmFic, SoFic and HYPE constructs. For structure determination, VbhA and VbhT(FIC) were co-expressed in the BL21(DE3) strain of *E. coli*, and NmFic, SoFic(Δ 8) and NmFic(SE/AA) in the BL21-AI strain of *E. coli*. VbhA-VbhT(FIC) and NmFic were purified by affinity chromatography followed by size exclusion chromatography. An additional anion exchange chromatography step was performed for NmFic(Δ 8) and NmFic(SE/AA). Diffraction data were collected at beamline X06SA of the Swiss Light Source. For NmFic(Δ 8) and NmFic(SE/AA), we obtained phases by molecular replacement using the NmFic structure (PDB code 2G03) as a search model. The VbhA-VbhT(FIC) complex was solved by molecular replacement using a fragment of BepA (PDB code 2JK8).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 July; accepted 24 November 2011.

Published online 22 January 2012.

1. Worby, C. A. *et al.* The fic domain: regulation of cell signaling by adenylation. *Mol. Cell* **34**, 93–103 (2009).

2. Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* **323**, 269–272 (2009).
3. Roy, C. R. & Mukherjee, S. Bacterial FIC proteins AMP up infection. *Sci. Signal.* **2**, pe14 (2009).
4. Mattoo, S. *et al.* Comparative analysis of *Histophilus somni* immunoglobulin-binding protein A (IbpA) with other Fic domain-containing enzymes reveals differences in substrate and nucleotide specificities. *J. Biol. Chem.* **286**, 32834–32842 (2011).
5. Luong, P. *et al.* Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J. Biol. Chem.* **285**, 20155–20163 (2010).
6. Xiao, J., Worby, C. A., Mattoo, S., Sankaran, B. & Dixon, J. E. Structural basis of Fic-mediated adenylation. *Nature Struct. Mol. Biol.* **17**, 1004–1010 (2010).
7. Dehio, C. *et al.* *Bartonella schoenbuchii* sp. nov., isolated from the blood of wild roe deer. *Int. J. Syst. Evol. Microbiol.* **51**, 1557–1565 (2001).
8. Engel, P. *et al.* Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella*. *PLoS Genet.* **7**, e1001296 (2011).
9. Schulein, R. *et al.* A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. *Proc. Natl Acad. Sci. USA* **102**, 856–861 (2005).
10. Palanivelu, D. V. *et al.* Fic domain-catalyzed adenylation: insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Sci.* **20**, 492–499 (2011).
11. Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M. & Komano, T. Involvement of cyclic AMP and its receptor protein in filamentation of an *Escherichia coli* fic mutant. *J. Bacteriol.* **151**, 807–812 (1982).
12. Engelberg-Kulka, H., Amitai, S., Kolodkin-Gal, I. & Hazan, R. Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genet.* **2**, e135 (2006).
13. Garcia-Pino, A. *et al.* Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *J. Biol. Chem.* **283**, 30821–30827 (2008).
14. Kinch, L. N., Yarbrough, M. L., Orth, K. & Grishin, N. V. Fido, a novel AMPylation domain common to Fic, Doc, and AvrB. *PLoS ONE* **4**, e5818 (2009).
15. Das, D. *et al.* Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gjl24375750) from *Shewanella oneidensis* MR-1 at 1.6 Å resolution. *Proteins* **75**, 264–271 (2009).
16. Mukherjee, S. *et al.* Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103–106 (2011).
17. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Glatter for mass spectrometry analysis of samples at the Core Proteomics facility. We thank the staff of beamline X06SA of the Swiss Light Source for assistance with data acquisition. We are grateful to G. Pluschke for providing the genomic DNA of *Neisseria meningitidis*, the ASU Biodesign Institute for providing the plasmid enclosing the *Shewanella oneidensis* Fic protein and S. Mattoo and J. Dixon for providing the pET-GSTX plasmids enclosing HYPE and HYPE(H295A). We also thank D. Bumann and A. Boehm for providing plasmid pC10E and *E. coli* strain AB472, respectively. This work was supported by grants 3100-061777 and 3100-138414 from the Swiss National Science Foundation (to C.D. and T.S., respectively), and grant 51RT O_126008 (InfecX) in the frame of the SystemsX.ch Swiss Initiative for Systems Biology (to C.D.).

Author Contributions P.E., F.V.S. and A.H. cloned recombinant plasmids. P.E. discovered and physiologically characterized VbhT-VbhA as a toxin-antitoxin module and carried out the bioinformatic analysis. A.G. expressed, purified and crystallized VbhA-VbhT(FIC), NmFic(SE/AA) and NmFic(Δ 8), and determined their structures. F.V.S. expressed, purified and crystallized NmFic with AMPPNP and determined the structure. A.G. and A.H. performed the adenylation assays. A.H. carried out the growth curve experiments. A.S. conducted the mass spectrometry analysis. All authors contributed to experimental design and data analysis. The manuscript was written by P.E., A.G., T.S. and C.D.

Author Information The atomic coordinates of VbhA-VbhT(FIC) and the complexes of NmFic, NmFic(SE/AA) and NmFic(Δ 8) with AMPPNP have been deposited in the Protein Data Bank under accession codes 3SHG, 3S6A, 3SN9 and 3SE5, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.D. (Christoph.Dehio@unibas.ch) or T.S. (Tilman.Schirmer@unibas.ch).

METHODS

Identification of VbhA homologues. VbhA of *B. schoenbuchensis* was queried against a database composed of translated open reading frames (>10 amino acids) identified in the 500-bp upstream region of all *fic* loci (PFAM release 24, 2,189 proteins). Nine rounds of Psi-Blast were performed and hits with an *E*-value <1 manually validated.

***E. coli* toxicity tests and cell filamentation.** *E. coli* AB472, a derivative of MG1655, was transformed with VbhT-expressing plasmids and always handled in LB containing 1% glucose. A single colony was picked, resuspended in 20 µl of LB and plated on LB plates containing 100 µM IPTG (induction) or 1% glucose (repression). Plates were incubated overnight at 37 °C. Growth curves were acquired by measuring optical density of liquid cultures in LB containing 100 µM IPTG that had been inoculated from overnight cultures of single colonies and were grown at 30 °C with continuous shaking. Cell filamentation was visualized by co-transformation of plasmid pC10E that constitutively expresses GFP and examined using fluorescence microscopy.

Conjugation experiments for VbhT toxicity tests in *Bartonella*. Plasmids were introduced into *Bartonella* strains by conjugation from *E. coli* using three-parental mating. *Bartonella* strains were grown for 36–48 h at 35 °C with 5% CO₂ on Columbia base agar plates supplemented with 5% defibrinated sheep blood and 100 µg ml⁻¹ streptomycin. *E. coli* β2150 that harbours helper plasmid pRK2013, and *E. coli* β2150 that contains the VbhT-expressing plasmid were grown overnight at 37 °C in LB supplemented with 50 µg ml⁻¹ kanamycin or 30 µg ml⁻¹ chloramphenicol, respectively, and both media were also supplemented with diamminopimelic acid (DAP) and 1% glucose. After 16 h of incubation, *E. coli* strains were diluted 1:50 in fresh LB medium and grown to a optical density (OD) at 595 nm (OD_{595nm}) of 0.4–0.8. Subsequently, each *E. coli* strain was diluted to OD_{595nm} of 0.25, washed once and resuspended in supplemented M199 medium (containing 10% FCS and 1% glucose). *Bartonella* strains were collected in 1 ml and resuspended in 60 µl of M199 (OD_{595nm} = 1). Each *Bartonella* suspension was mixed with 20 µl of *E. coli* β2150 that harbours pRK2013 and 20 µl of *E. coli* β2150 that harbours the VbhT-expressing plasmid. The conjugation mix was distributed on a conjugation filter on a columbia blood agar (CBA) plate supplemented with 150 µl DAP and 150 µl 1% glucose. After 6 h of incubation under *Bartonella* growth conditions as described before, the bacteria were washed off the filter with 1 ml supplemented M199. Dilutions of bacterial suspensions were plated on lysogeny broth agar (LA) supplemented with DAP, 1% glucose and 30 µg ml⁻¹ chloramphenicol for selecting donors, on CBA supplemented with 1.2 µg ml⁻¹ chloramphenicol for selecting transconjugants, and on CBA supplemented with 100 µg ml⁻¹ streptomycin for selecting recipients. Agar plates were incubated under *Bartonella* growth conditions and colony-forming units were counted after 1 day for donors and after 7 days for recipients and transconjugants.

In vitro adenylation assay. Adenylation activity of VbhT was assessed in an assay using lysates of ectopically expressing *E. coli*. Bacterial pellets were resuspended in reaction buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1 mM EGTA, 15 mM MgCl₂, 140 µg ml⁻¹ RNase A and protease inhibitor cocktail (Roche)) and lysed by sonication. After clearing lysates by centrifugation, supernatants were used for experimentation or stored at -20 °C.

Adenylation reactions were prepared by supplementing 15 µl supernatant from expression cultures with 10 µCi [α -³²P]ATP (Hartmann Analytic) and 25 µl blank *E. coli* supernatant. Adenylation activity of NmFic and SoFic constructs was assessed by incubating 60 µg purified protein with 10 µCi [α -³²P]ATP (Hartmann Analytic) and 25 µl blank *E. coli* supernatant. Reactions were incubated for 1 h at 30 °C, resolved by SDS-PAGE, and adenylation was probed by autoradiography. For HYPE(FIC) assays, 15 ng of pure protein was incubated with 1.6 µg of purified GTPases.

Protein expression and purification. pFVS0011 vector (encoding VbhA and VbhT(FIC)) was transformed into *E. coli* BL21 (DE3). *E. coli* cultures were grown at 37 °C in LB medium supplemented with 50 µg ml⁻¹ of kanamycin to an OD_{595nm} of 0.6 before induction with 0.3 mM IPTG for 16 h at 23 °C. Vectors pFVS0015 (carrying the *NmFic* gene), pFVS0016 (encoding NmFic(Δ8)), pFVS0037 (encoding NmFic(SE/AA)) were transformed into BL21-AI cells. Cells were incubated overnight in 750 ml LB medium that was supplemented with 50 µg ml⁻¹ kanamycin and 1% glucose at 23 °C at 200 r.p.m. until OD_{595nm} of 2 was reached. Bacterial pellets were resuspended in terrific broth medium containing 50 µg ml⁻¹ kanamycin to obtain an OD_{595nm} of approximately 1. Protein expression was induced at 23 °C with 0.1% arabinose and 0.5 mM IPTG for 23 h at 200 r.p.m. Plasmids harbouring HYPE(FIC) and SoFic constructs were transformed in *E. coli* Rosetta (DE3) cells and BL21-AI cells, respectively. The proteins were expressed as described for NmFic.

Cells containing overexpressed VbhA-VbhT(FIC) and NmFic were resuspended in lysis buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl and 25 mM imidazole and disrupted using French press. Cell debris was pelleted by

ultracentrifugation and the supernatant was applied to a His-Trap column (GE Healthcare). The stable complex was eluted with a gradient of elution buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl and 500 mM imidazole. The protein was then concentrated and injected on a Superdex 75 16/60 gel filtration column (GE Healthcare) equilibrated with 10 mM Tris (pH 7.5) and 100 mM NaCl. The pure proteins were concentrated to 6 mg ml⁻¹ for VbhA-VbhT(FIC) and 53 mg ml⁻¹ for NmFic.

The same purification protocol was used for NmFic(Δ8) and NmFic (SE/AA), but with an additional intermediate purification step. After affinity purification, the proteins were adjusted to 20 mM Tris (pH 8.5), 25 mM NaCl, applied to a Resource-Q anion exchange column (Amersham Biosciences) and eluted with a linear gradient of NaCl. Peak fractions were concentrated and further purified by gel filtration chromatography. Purified proteins in 10 mM Tris (pH 7.8), 100 mM NaCl were concentrated to 30 mg ml⁻¹ for NmFic(Δ8) and 51 mg ml⁻¹ for NmFic(SE/AA). SoFic and SoFic(E73G) were purified as described previously¹⁵. HYPE(FIC), HYPE(FIC/H295A) and HYPE(FIC/E234G) were purified in the same way as NmFic. GST-tagged Cdc42 and Rac1 were expressed and purified as described previously^{18,19}.

Crystallization. All crystals were obtained at 20 °C (except for NmFic crystals, which were obtained at 4 °C) using the hanging-drop vapour diffusion method after mixing 1 µl protein solution with 1 µl reservoir solution. VbhA-VbhT(FIC) and NmFic(SE/AA) crystallized in 23% (w/v) PEG 3350 and 0.2 M di-ammonium tartrate, and were cryoprotected with 25% (w/v) PEG 3350, 0.2 M di-ammonium tartrate and 10% glycerol. NmFic crystallized in 5% 2-propanol, 0.1 M MES pH 6.0 and 0.1 M Ca-acetate, and was transferred into 8% 2-propanol, 0.1 M MES pH 6.0, 0.1 M Mg-acetate and 15% glycerol, then into 30% glycerol for cryoprotection. NmFic(Δ8) crystallized with 44% (v/v) PEG 600, 0.1 M Na-citrate pH 5.6. No cryoprotection was needed for data collection. In each case, the substrate analogue complex was produced by crystal soaking for 2 h with 10 mM AMPPNP, 10 mM MgCl₂.

Structure determination. Statistics of data collection and refinement are given in Supplementary Table 1. Diffraction data were collected at beamline X06SA (PXIII) of the Swiss Light Source (λ = 1.0 Å) at 100 K on a MAR CCD detector, processed using MOSFLM²⁰, and scaled with SCALA²¹. The structures were determined by molecular replacement (PHASER²²) using a BepA fragment (PDB code 2JK8¹⁰, residues 30–194) or the uncomplexed NmFic structure (PDB code 2G03, unpublished, Midwest Center for Structural Genomics) as models for structure solution of VbhA-VbhT(FIC) and the different NmFic proteins, respectively. A structure solution of wild-type and mutant NmFic in complex with AMPPNP was straightforward, whereas a weak solution (RFZ = 5.1, TFZ = 3.6) with poor phasing power was obtained for VbhA-VbhT(FIC). The partial model lacking VbhA was refined by rigid-body refinement using REFMAC5 (ref. 23) with three bodies to an *R*_{free} of 52%. Model extension using the module AutoBuild of the PHENIX package²⁴ yielded an almost complete model (*R*/*R*_{free} = 30.8%/35.2%). The remainder of the molecule was traced manually with COOT²⁵ and then by full refinement using PHENIX²⁴. The Ramachandran plot showed that more than 99% of the residues are in favoured regions of the four structures. The figures were generated with Dino (<http://www.dino3d.org>).

Prediction of inhibition motif in Fic proteins. All Fic proteins (PFAM release 24, 2,189 proteins) were subjected to a profile-to-profile comparison with sequences from the PDB using HHpred¹⁷. HHpred builds an alignment of homologues for each query sequence by using iterations of PSI-BLAST searches against the non-redundant database. Secondary structures are then predicted on the PSI-BLAST alignment using PSIPRED²⁶. On the basis of this data, a profile Hidden Markov Model (HMM) is generated. Each query profile is compared with the pre-computed HMMs of the proteins in the PDB to identify structural homologues. In terms of query profiles, the PDB profiles include secondary structure information derived from their three-dimensional structure. We analysed the pairwise profile alignments of each Fic protein with the eight different Fic family members deposited in the PDB. A Fic protein was predicted to belong to class II or class III if the templates' inhibitory motifs were aligned to a corresponding query sequence in the profile alignments (see Supplementary Information, section 1).

Phylogenetic analysis of Fic proteins. Phylogenetic trees of the Fic family were inferred with FastTree 2 (ref. 27) and RAXML 7.0.4 (ref. 28). Trees were built on the amino acid alignment provided by the PFAM database. Unaligned overhanging ends were trimmed off and identical sequences were reduced to one representative. We used local support values based on the Shimodaira-Hasegawa test to estimate the reliability of the tree inferred with FastTree2. The RAXML tree was inferred using the PROTMIXWAGF model and 25 rate categories.

Liquid chromatography-mass spectrometry analysis. 2.5 µM of purified NmFic, NmFic(Δ8) or NmFic(SE/AA) were incubated in reaction buffer (10 mM Tris, pH 8, 100 mM NaCl) in the presence or absence of 50 µM ATP and 50 µM MgCl₂ for 1 h. Proteins were reduced in 5 mM TCEP, alkylated in

10 mM iodoacetamid and digested with sequencing grade trypsin (Promega). The generated peptides were purified with C18 Microspin columns (Harvard Apparatus) and analysed using liquid chromatography-mass spectrometry (LC-MS) or MS on an easy nano-LC system coupled to an LTQ-Orbitrap-Velos mass spectrometer (both from Thermo-Fisher Scientific), as recently described²⁹ using a linear gradient from 95% solvent A (0.15% formic acid, 2% acetonitrile) and 5% solvent B (98% acetonitrile, 0.15% formic acid) to 35% solvent B over 40 min. The data acquisition mode was set to obtain one high-resolution MS scan in the Fourier Transform (FT) part of the mass spectrometer at a resolution of 60,000 (full width at half maximum) and MS-MS scans in the linear ion trap of the 20 most intense ions. The resulting MS2 scans were searched against a *N. meningitidis* protein database containing the target protein sequence, including NmFic and NmFic(SE/AA) sequences, that was obtained from EBI (<http://www.ebi.ac.uk>) using the SEQUEST search algorithm provided in the Proteome Discoverer software package (Thermo-Fisher Scientific). *In silico* trypsin digestion was performed after lysine and arginine (unless followed by proline), with a tolerance of two missed cleavages in fully tryptic peptides. Database search parameters were set to allow phosphoadenosine modification (+329.05252 Da) of threonine and tyrosine residues as variable modification and carboxyamidomethylation (+57.021464 Da) of cysteine residues as fixed modification. The fragment mass tolerance was set to 0.8 Da and the precursor mass tolerance to 15 p.p.m.

Strain construction. For toxicity experiments, the *vbbT* wild-type gene (FN645515) from *B. schoenbuchensis* R1 was cloned into pRSF-Duet1 (pPE0017, His₆-tagged *vbbT*). *VbbT*(H136A) (pPE0034) was constructed by introducing a two-base-pair point mutation in the FIC motif of *vbbT* of pPE0017, as described elsewhere³⁰. Plasmid co-expressing *VbbT* and *VbhA* (*VbbT/VbhA*) was constructed by cloning *vbhA* (FN645515) amplified from *B. schoenbuchensis* R1 into pRSF-Duet1 (pPE0020, HA-tagged *vbhA*). *vbbT* was then cloned into pPE0020, resulting in pPE0021. To construct *VbbT*-expressing plasmids for *Bartonella*, *vbbT* from *B. schoenbuchensis* R1 was cloned into vector pMMB206 (ref. 31) (pVbbT, HA-tagged *vbbT*). pVbbT(H136A) was constructed from pVbbT as described before.

The in-frame deletion of the complete *vbhA/vbbT* operon in *B. schoenbuchensis* (*Bsch ΔvbbA/vbbT*) was generated as described previously by a two-step gene replacement procedure⁹. The mutagenesis vector pPE3005 was constructed by ligating a cassette with the flanking regions of the in-frame deletion into pTR1000⁹.

For protein purification, the full-length *vbhA* gene and part of the *vbbT* gene encoding the FIC domain (*vbbT*(FIC)), amino acid residues 1–198, His₆-tagged) were amplified from plasmid pPE0021 and cloned into the pRSF-Duet1 vector (pFVS0011). *VbhA*(E24G)/*VbbT*(FIC) expression plasmid (pFVS0065) was generated by introducing a two-base-pair mutation in pFVS0011. The *NmFic* gene was amplified with an N-terminal His₆-tag from *N. meningitidis* from coding region of amino acid residues 11–191 and from coding region of amino acid residues 11–167 to generate plasmids expressing NmFic (pFVS0015) and NmFic(Δ8) (pFVS0016), respectively. The S182A/E186A double mutant construct (NmFic(SE/AA)) was generated by introducing two subsequent point mutations in pFVS0015. The E186G mutant construct (NmFic(E186G)) was

generated by the same approach. The *SoFic* gene was amplified from plasmid (ASU biodesign institute, Clone ID SoCD00104192) and cloned with an N-terminal His₆-tag into pRSF-Duet1 (pFVS0040). The *SoFic*(E73G) plasmid (pFVS0058) was generated by introducing a two-base-pair point mutation in pFVS0040. GST-HYPE(E234G) (pFVS0064) was generated by introducing point mutations in the plasmid containing GST-HYPE. From these plasmids, shorter constructs (HYPE(FIC), HYPE(FIC/E234G) and HYPE(FIC/H295A)) only carrying the FIC domain of HYPE (from amino acids 187 to 437) were generated. For the expression of human Cdc42 and Rac1, the *Cdc42-Q61L* and *Rac1-Q61L* coding sequences were amplified from plasmid pRK5myc L61 Cdc42³² and pRK5FLAG L61 Rac1 (ref. 32) and cloned into pGex6p1 with an N-terminal GST-tag, resulting in pAH088 and pAH060, respectively. The wild-type variants of Cdc42 and Rac1 were generated from the mutant constructs through polymerase chain reaction (PCR)-based site-directed mutagenesis³⁰ (resulting in pAH059 and pAH071). All primers and the resulting vectors are summarized in Supplementary Tables 5 and 6.

18. Self, A. J. & Hall, A. Purification of recombinant Rho/Rac/G25K from *Escherichia coli*. *Methods Enzymol.* **256**, 3–10 (1995).
19. Smith, S. J. & Ritinger, K. Preparation of GTPases for structural and biophysical analysis. *Methods Mol. Biol.* **189**, 13–24 (2002).
20. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62**, 48–57 (2006).
21. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
22. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
23. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
24. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
25. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
26. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
27. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
28. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
29. Schmidt, A. *et al.* Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* **7**, 510 (2011).
30. Zheng, L., Baumann, U. & Reymond, J. L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
31. Dehio, C. & Meyer, M. Maintenance of broad-host-range incompatibility group P and group Q plasmids and transposition of Tn5 in *Bartonella henselae* following conjugal plasmid transfer from *Escherichia coli*. *J. Bacteriol.* **179**, 538–540 (1997).
32. Rhomberg, T. A., Truttmann, M. C., Guye, P., Ellner, Y. & Dehio, C. A translocated protein of *Bartonella henselae* interferes with endocytic uptake of individual bacteria and triggers uptake of large bacterial aggregates via the invasome. *Cell. Microbiol.* **11**, 927–945 (2009).

Cysteinyl leukotriene type I receptor desensitization sustains Ca^{2+} -dependent gene expression

Siaw-Wei Ng¹, Daniel Bakowski¹, Charmaine Nelson¹, Ravi Mehta¹, Robert Almeyda², Grant Bates² & Anant B. Parekh¹

Receptor desensitization is a universal mechanism to turn off a biological response; in this process, the ability of a physiological trigger to activate a cell is lost despite the continued presence of the stimulus. Receptor desensitization of G-protein-coupled receptors involves uncoupling of the receptor from its G-protein or second-messenger pathway followed by receptor internalization¹. G-protein-coupled cysteinyl leukotriene type I (CysLT1) receptors regulate immune-cell function and CysLT1 receptors are an established therapeutic target for allergies, including asthma². Desensitization of CysLT1 receptors arises predominantly from protein-kinase-C-dependent phosphorylation of three serine residues in the receptor carboxy terminus³. Physiological concentrations of the receptor agonist leukotriene C_4 (LTC_4) evoke repetitive cytoplasmic Ca^{2+} oscillations, reflecting regenerative Ca^{2+} release from stores, which is sustained by Ca^{2+} entry through store-operated calcium-release-activated calcium (CRAC) channels⁴. CRAC channels are tightly linked to expression of the transcription factor *c-fos*⁵, a regulator of numerous genes important to cell growth and development⁶. Here we show that abolishing leukotriene receptor desensitization suppresses agonist-driven gene expression in a rat cell line. Mechanistically, stimulation of non-desensitizing receptors evoked prolonged inositol-trisphosphate-mediated Ca^{2+} release, which led to accelerated Ca^{2+} -dependent slow inactivation of CRAC channels and a subsequent loss of excitation–transcription coupling. Hence, rather than serving to turn off a biological response, reversible desensitization of a Ca^{2+} mobilizing receptor acts as an ‘on’ switch, sustaining long-term signalling in the immune system.

The paradox of receptor desensitization is how long-term responses can be evoked if the receptor is inactivated. This is a particularly acute problem in immune cells, in which cell differentiation and clonal selection develop over hours in the continued presence of external cues.

Stimulation of rat basophilic leukaemia (RBL-1) cells with LTC_4 , acting exclusively on CysLT1 receptors^{7,8} (Supplementary Fig. 1), led to cytoplasmic Ca^{2+} signals (Fig. 1a) followed by robust expression of *c-fos* at both messenger RNA (Fig. 1b, c)⁴ and protein levels (Fig. 1d, e). Maximal activation of CRAC channels with thapsigargin led to a similar increase in *c-fos* expression (Fig. 1b–e). Both LTC_4 and thapsigargin induce *c-fos* expression through the build-up of Ca^{2+} microdomains near open store-operated CRAC channels^{4,5}. Thapsigargin led to a larger, more sustained Ca^{2+} signal than LTC_4 (Fig. 1a)^{4,5} and the rate of Ca^{2+} entry through CRAC channels was approximately twofold more for thapsigargin than LTC_4 (Fig. 1f), consistent with patch-clamp recordings⁷. The similar increase in Ca^{2+} -dependent *c-fos* expression to LTC_4 and thapsigargin was therefore surprising, given the significant difference in CRAC channel activation.

We considered various explanations for why CysLT1 receptor activation and thapsigargin evoked similar *c-fos* levels despite marked differences in the extent of CRAC channel activation. These included: (1) CysLT1 receptors tapped into a different signalling mechanism linking CRAC channel microdomains to *c-fos* expression; (2) local Ca^{2+} entry through CRAC channels was larger after receptor activation

because LTC_4 hyperpolarized the membrane potential; and (3) cytoplasmic Ca^{2+} and protein kinase C (PKC) interacted synergistically to drive gene expression in response to CysLT1 receptor activation. Evidence against these possibilities is presented in Supplementary Figs 2–4. Instead, gene expression showed high sensitivity to Ca^{2+} entry, enabling CysLT1 receptor activation to couple effectively to *c-fos* transcription (Supplementary Fig. 5), as well as rapidly and high gain. Combined, this ensures efficient gene expression to bursts of CRAC channel activity after physiological levels of receptor stimulation.

Experiments described in Fig. 1g and h revealed an important role for PKC in receptor-dependent gene expression. The structurally distinct PKC blockers G06983 and calphostin C abolished *c-fos* expression (Fig. 1g, h). PKC block had a marked effect on the Ca^{2+} signal evoked by agonist. Whereas cytoplasmic Ca^{2+} oscillations were routinely observed with LTC_4 (Fig. 1i), the response was converted into a large, single, slowly decaying Ca^{2+} spike after PKC inhibition (Fig. 1i, j). Acute stimulation with PMA in the absence of LTC_4 failed to induce significant *c-fos* expression (data not shown)⁹, demonstrating that PKC activity per se was not sufficient to induce *c-fos* expression in these cells. The PKC inhibitors had no effect on thapsigargin-evoked *c-fos* expression (Fig. 1k) or cytoplasmic Ca^{2+} signals (Fig. 1l). Thapsigargin (2 μM) activates CRAC channels maximally (Supplementary Fig. 5) and, by blocking SERCA pumps (which can be located near CRAC channels¹⁰), reduces the decay of Ca^{2+} gradients radiating from the plasma membrane. It is possible that other non-receptor-dependent stimuli, which raise local Ca^{2+} levels less effectively than 2 μM thapsigargin, might activate *c-fos* in a manner dependent on basal PKC activity, but this activity would be unusual in that it is not stimulated acutely by PMA in the presence of sub-maximal CRAC channel activation (Supplementary Fig. 4).

In RBL cells, exposure to the phorbol ester PMA for several hours downregulates several PKC isozymes¹¹. Using this protocol, we found that *c-fos* expression was substantially reduced in response to CysLT1 receptor stimulation (Fig. 1m), whereas no significant reduction was seen when thapsigargin was used instead (Fig. 1m) or when inactive 4 α -phorbol replaced PMA (data not shown). Similar to PKC blockers, the Ca^{2+} signal to LTC_4 was prolonged after PKC downregulation (Fig. 1n). This prolonged Ca^{2+} signal did not reflect a change in Ca^{2+} clearance mechanisms (Supplementary Fig. 6); instead, it is characteristic of loss of receptor desensitization, particularly for CysLT1 receptors, in which desensitization is mediated predominantly by PKC³ and prevention of desensitization leads to broader Ca^{2+} signals¹². Inhibition of CysLT1 receptor desensitization is predicted to lead to greater inositol trisphosphate (InsP_3) production and hence more extensive Ca^{2+} store emptying. Several findings are consistent with this. First, Ca^{2+} release to LTC_4 lasted approximately five times longer when PKC was blocked than in control cells (Fig. 2a, expanded in inset). Second, the amount of Ca^{2+} remaining within the stores, measured as the ionomycin-sensitive Ca^{2+} response¹³, was substantially less after activation of CysLT1 receptors in the presence of PKC block than in control cells (Fig. 2a). Third, InsP_3 production, measured using the green fluorescent protein–pleckstrin homology

¹Department of Physiology, Anatomy and Genetics, University of Oxford, Parks Road, Oxford OX1 3PT, UK. ²Department of Ear-Nose-Throat Surgery, John Radcliffe Hospital, Oxford OX3 9DH, UK.

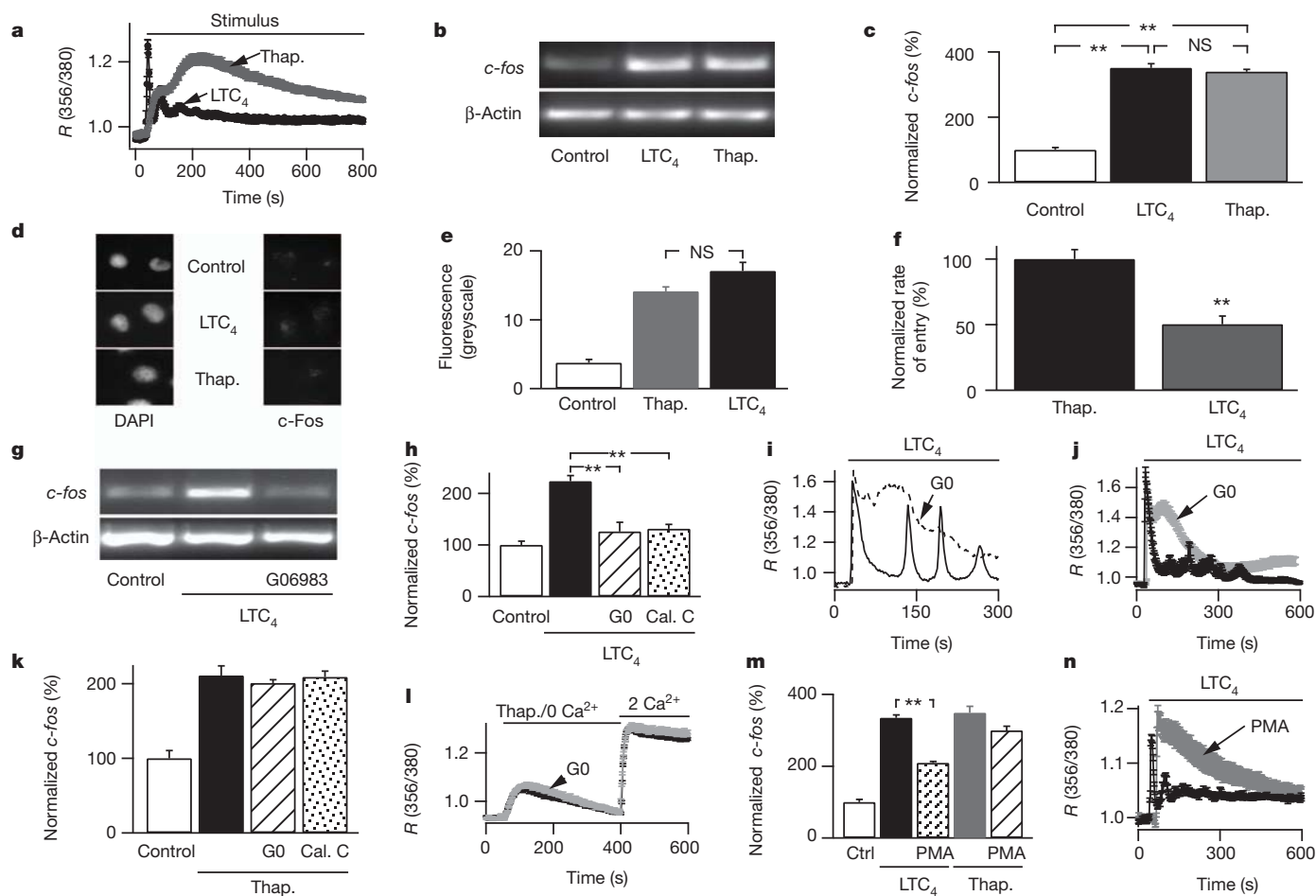


Figure 1 | CysLT1-receptor-dependent *c-fos* expression requires PKC.

a, Averaged Ca^{2+} signals to LTC_4 and thapsigargin (Thap.) are compared (>50 cells per graph). R (356/380), ratio of fluorescence measured at 356 and 380 nm. **b**, *c-fos* expression is compared between control (non-stimulated), 160 nM LTC_4 - and 2 μM thapsigargin-stimulated cells. Stimulus was present for 8 min. **c**, Histograms show averaged responses from three independent experiments. LTC_4 and thapsigargin groups were different from control ($P < 0.001$), but not from one another ($P > 0.3$; analysis of variance (ANOVA)). **d**, Cells stained with antibody against *c-Fos* protein. DAPI, 4',6-diamidino-2-phenylindole. **e**, Aggregate data are compared ($n > 20$ per bar). Thapsigargin and LTC_4 groups were different from control ($P < 0.001$) but not from one another ($P = 0.11$). **f**, Ca^{2+} entry rate was measured after readmission of Ca^{2+} to cells stimulated with LTC_4 or thapsigargin in Ca^{2+} -free solution ($*P < 0.01$). **g**, G06983 (1 μM ; 10 min pre-treatment) suppresses LTC_4 -induced *c-fos* expression. **h**, Histogram comparing the effects of PKC blockers. LTC_4 control group (LTC_4 in the absence of PKC block) was different from the other groups

($P < 0.01$). There were no significant differences between the other groups. G0, G06983. Cal. C, calphostin C. **i**, Single-cell Ca^{2+} signals to LTC_4 are compared for the conditions shown. **j**, Averaged data are compared (>45 cells for each condition). **k**, Histogram showing *c-fos* expression to thapsigargin in the presence of PKC blockers. All thapsigargin-treated groups were significantly different from control ($P < 0.001$) but were not significantly different from one another. **l**, Ca^{2+} signals to thapsigargin are unaffected by PKC block. 0 Ca^{2+} , Ca^{2+} -free external solution; 2 Ca^{2+} , external solution containing 2 mM Ca^{2+} . **m**, Downregulation of PKC (PMA; 500 nM, 24 h) reduces LTC_4 - but not thapsigargin-induced *c-fos* expression (data from four independent experiments). All stimulated groups were significantly different from control ($P < 0.01$). For LTC_4 the PMA group was different from the LTC_4 control ($P < 0.01$). For thapsigargin, the PMA groups were not different from the thapsigargin control ($P = 0.07$). Ctrl, control. **n**, PKC downregulation alters the LTC_4 -evoked Ca^{2+} signal. Error bars show data \pm s.e.m. NS, not significant.

(GFP-PH) construct¹⁴, increased to a greater extent when PKC was inhibited (Fig. 2b).

Cytoplasmic Ca^{2+} inhibits CRAC channels through mechanisms of fast and slow inactivation¹⁵. The prolonged Ca^{2+} release evoked by LTC_4 in the presence of non-desensitizing receptors could therefore inactivate CRAC channels to suppress agonist-evoked gene expression. In support of this, accumulation of the slow Ca^{2+} chelator EGTA in the cytoplasm rescued gene expression to CysLT1 receptor activation in the presence of PKC block (Fig. 2c, d). Ca^{2+} -dependent fast inactivation of CRAC channels is unlikely to contribute here because (1) it is unaffected by the slow chelator EGTA^{16,17}, which reversed the inhibitory effects of PKC block (Fig. 2c, d), and (2) the rate and extent of fast inactivation were unaltered by CysLT1 receptor activation in the presence of PKC downregulation (Fig. 3a). Instead, Ca^{2+} -dependent slow inactivation is likely to be the dominant mechanism because (1) it too is suppressed by cytoplasmic EGTA^{18,19}, (2) the

Ca^{2+} dependence of slow inactivation has a dissociation constant (K_d) of $\sim 0.5 \mu\text{M}$ and full block occurs at $\sim 1 \mu\text{M}$ (Fig. 3b), which is similar to the peak Ca^{2+} rise evoked by LTC_4 in the presence of PKC inhibitors or following downregulation of PKC ($0.87 \pm 0.1 \mu\text{M}$), and (3) Ca^{2+} -dependent slow inactivation develops with a time course similar to the duration of the prolonged Ca^{2+} rise caused by LTC_4 following loss of PKC activity^{18,19}.

If prolonged Ca^{2+} release to non-desensitizing CysLT1 receptors leads to slow inactivation of CRAC channels, then development of a calcium-release-activated calcium current (I_{CRAC}) to a subsequent stimulus should be impaired. Pre-activation of CysLT1 receptors reduced I_{CRAC} evoked by thapsigargin but only in the presence of PKC block (Fig. 3c). No such inhibitory effect was seen when cells were dialysed with a strongly buffered Ca^{2+} -containing pipette solution, which prevents the development of slow inactivation (Fig. 3d). Increasing the time between the termination of Ca^{2+} release and

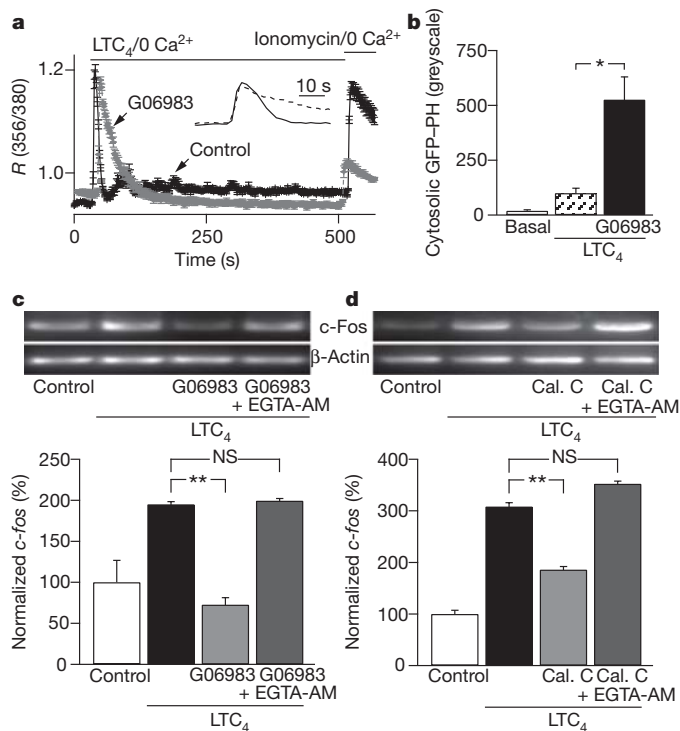


Figure 2 | Gene expression to non-desensitizing CysLT1 receptors is rescued by preventing a cytoplasmic Ca^{2+} rise. **a**, Stimulation with LTC_4 in the presence of G06983 evokes a more sustained Ca^{2+} release response, and this leads to more extensive store depletion (measured through the extent of Ca^{2+} release evoked by $5 \mu\text{M}$ ionomycin). Both LTC_4 and ionomycin were applied in Ca^{2+} -free external solution. Inset compares the kinetics of Ca^{2+} release. **b**, Cytosolic GFP-PH levels, a measure of InsP_3 levels, rise when CysLT1 receptors are stimulated in the presence of G06983. **c**, Top, loading cells with the Ca^{2+} chelator EGTA prevents loss of gene expression to agonist when PKC is blocked. Bottom, aggregate data from five independent gels are summarized. **d**, As in **c**, but calphostin C was used to block PKC instead. Bottom, aggregate data from three independent gels are summarized. Error bars show data \pm s.e.m. NS, not significant.

subsequent store-operated Ca^{2+} entry should enable some recovery from Ca^{2+} -dependent slow inactivation and this should partially rescue gene expression. When Ca^{2+} influx was evoked a few minutes after Ca^{2+} release, significant, albeit incomplete, rescue of Ca^{2+} entry (Fig. 3e) and *c-fos* transcription (Fig. 3f) occurred in cells stimulated with LTC_4 in the presence of PKC inhibition. Hence, allowing CRAC channels time to recover from Ca^{2+} -dependent inactivation results in partial rescue of agonist-driven gene expression.

Our attempts to express the PKC-insensitive CysLT1 receptor, in which S313, S315 and S316 had been mutated to alanines, were thwarted by the difficulty of expressing these receptors³, although in a few cells we observed that Ca^{2+} oscillations to LTC_4 were less frequent (3.1 ± 0.5 versus 5.4 ± 0.4 , 4 and 6 cells, respectively) and the initial spike was a little broader (~ 1.25 -fold) than mock-transfected cells.

To place our findings in a physiological context, we turned to the human nasal polyp, which is rich in mast cells²⁰. The polyp and associated nasal mucosa are largely self-contained, providing an excellent 'quasi *in vivo* human system'. Mast cells from polyps, acutely isolated from patients undergoing surgery, respond to LTC_4 and express functional CRAC channels^{7,8}. Stimulation with LTC_4 activated *c-Fos* protein expression in mast cells isolated from polyps (Fig. 3g) and this was reduced by pre-treatment with either calphostin C or G06983 (Fig. 3h). PKC inhibitors had no inhibitory effect when thapsigargin was used instead.

Western blots revealed the presence of Ca^{2+} -dependent PKC- α , - β and - ζ isozymes^{11,21} but only faint expression of PKC- δ and - ϵ (Fig. 4a).

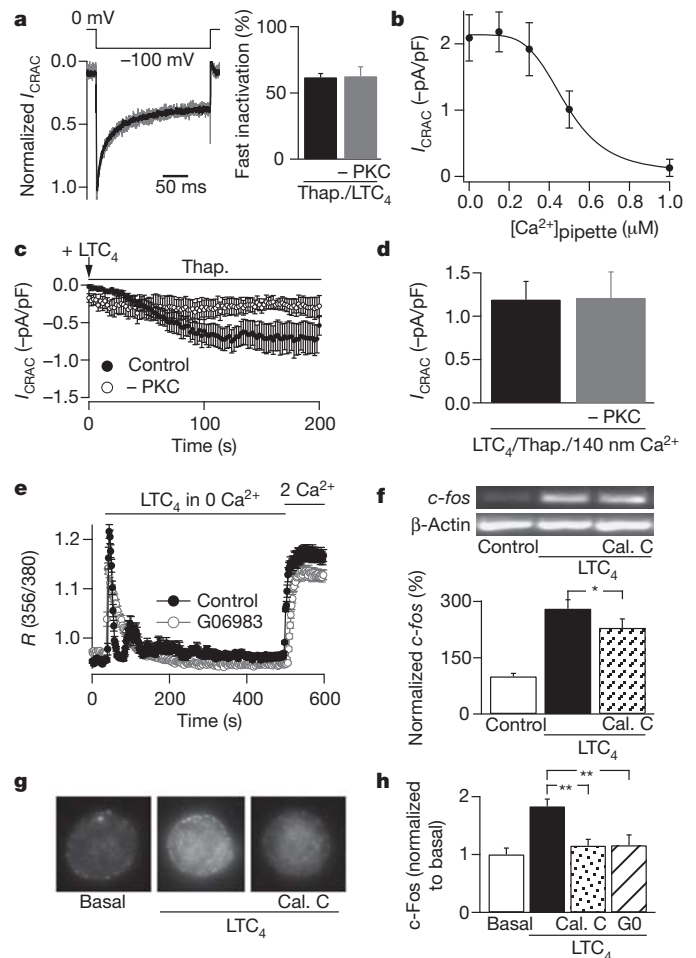


Figure 3 | Ca^{2+} -dependent slow inactivation underlies suppression of *c-Fos* expression to non-desensitizing CysLT1 receptors. **a**, Ca^{2+} -dependent fast inactivation is unaffected by non-desensitizing receptors (labelled -PKC). Cells were stimulated with LTC_4 (160 nM) before breaking in with a pipette solution containing thapsigargin and buffered Ca^{2+} (140 nM) and fast inactivation was measured within 60 s of break-in. **b**, Dependence of Ca^{2+} -dependent slow inactivation on patch pipette Ca^{2+} concentration. pA, picoamperes; pF, pF, pF. **c**, Stimulation of non-desensitizing receptors with LTC_4 before break-in significantly reduced the size of I_{CRAC} that developed in response to dialysis with thapsigargin in weak buffer (0.2 mM EGTA). **d**, As in **c**, but cells were dialysed with a pipette solution containing strong Ca^{2+} buffer (10 mM EGTA, 140 nM free Ca^{2+}). **e**, Store-operated Ca^{2+} entry recovers partially by increasing the time interval between Ca^{2+} release and subsequent Ca^{2+} entry. **f**, *c-fos* expression to non-desensitizing receptor stimulation is rescued partially when Ca^{2+} entry occurs several minutes after Ca^{2+} release has reached completion. **g**, *c-Fos* expression in human nasal mast cells after CysLT1 receptor activation is suppressed by PKC inhibition. **h**, Aggregate data are compared (12–17 cells per bar; three patients each). Error bars show data \pm s.e.m. NS, not significant.

Overnight PMA exposure significantly reduced PKC- α and - β expression, but not PKC- ζ (Fig. 4a, b). Their weak expression made PKC- δ and - ϵ difficult to quantify. Confocal microscope studies confirmed robust expression of PKC- α , - β and - ζ (Fig. 4c), with barely detectable levels of PKC- δ and - ϵ (data not shown). Overnight PMA exposure significantly reduced PKC- α and - β but not - ζ at the cellular level (Fig. 4d). Knockdown of PKC- α using a targeted short interfering RNA (siRNA) approach (Fig. 4e) resulted in a broadening of the first Ca^{2+} oscillation evoked by LTC_4 , indicative of less receptor desensitization, and fewer Ca^{2+} oscillations in each cell (Supplementary Fig. 7). Knockdown of PKC- β had a much weaker effect on the Ca^{2+} oscillations (Supplementary Fig. 7). Knockdown of PKC- α or PKC- α plus - β simultaneously, but not PKC- β alone, reduced LTC_4 -driven

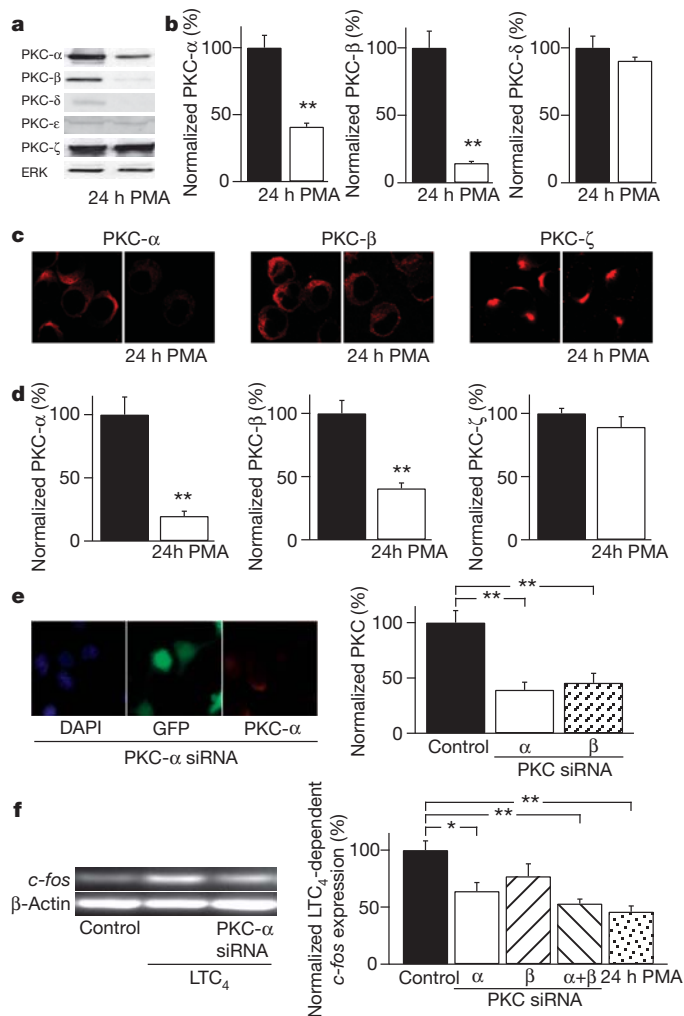


Figure 4 | PKC- α regulates CysLT1 receptor-driven *c-fos* transcription.

a, Expression of PKC- α , - β and - ζ (western blot) is shown in control cells and cells exposed to PMA for 24 h. **b**, Quantification of data from three independent experiments, as in **a**. **c**, Confocal microscope images of PKC expression for the conditions shown. Cells were fixed before analysis. Original magnification, $\times 100$. **d**, Quantification of images from experiments, as in **c**. **e**, siRNA against PKC- α or - β significantly reduces corresponding protein expression. Left, DAPI staining of nuclei; middle, GFP expression (indicating transfection); and right, PKC- α expression after siRNA-mediated knockdown. Right, aggregate data from four experiments are depicted. Both siRNA groups were different from control ($P < 0.005$). Error bars show data \pm s.e.m. NS, not significant. **f**, Knockdown of PKC- α , - β and PKC- α plus - β on LTC₄-dependent *c-fos* expression. Data are compared with mock-transfected cells. For comparison, 24 h exposure to PMA is included. All treated groups were significantly different from the LTC₄ control (black bar) group except siRNA PKC- β knockdown ($P > 0.1$). PKC- α plus - β and 24 h PMA groups had $P < 0.01$; PKC- α group had $P < 0.05$. Error bars show data \pm s.e.m. NS, not significant.

c-fos expression to an extent similar to that seen after overnight PMA treatment (Fig. 4f). Stimulation of non-G-protein-coupled FC ϵ RI antigenic receptors in RBL-2H3 cells activates *c-fos* expression primarily through PKC- δ and - ϵ ²². Although it is possible that these PKC isoforms also contribute to gene expression under our conditions, our results nevertheless suggest a major role for PKC- α in G-protein-coupled-receptor desensitization, and thus coupling to the nucleus.

Collectively, our findings reveal a counterintuitive function for desensitization of a phospholipase-C-coupled receptor. Rather than terminating a response, homologous receptor desensitization is essential for maintaining excitation–transcription coupling. Desensitization of CysLT1 receptors is mediated principally by PKC-dependent phosphorylation³. Prevention of receptor desensitization through either

acute block or degradation of PKC or after knockdown of PKC- α all led to loss of Ca²⁺-dependent gene expression, despite potentiation of Ca²⁺ release to agonist. Mechanistically, the prolonged Ca²⁺ release phase accelerated Ca²⁺-dependent slow inactivation of CRAC channels, resulting in loss of Ca²⁺ entry. Because Ca²⁺ microdomains near open CRAC channels drive *c-fos* expression, the decline in CRAC channel activity abolishes excitation–transcription coupling. The interval between Ca²⁺ oscillations after CysLT1 receptor activation is ~ 25 s (ref. 4). Because InsP₃ has a short half-life in the cytoplasm (~ 1 s)²³, receptor desensitization will presumably lower InsP₃ levels during the interspike interval. Store refilling will occur quickly and CRAC channel activity will be transient after CysLT1 receptor stimulation. The short duration of Ca²⁺ release and thus Ca²⁺ entry, determined by receptor desensitization, will ensure Ca²⁺-dependent slow inactivation does not develop, as this inhibitory mechanism requires a sustained Ca²⁺ rise for several seconds. It is therefore the kinetics of receptor desensitization and recovery from desensitization within a highly Ca²⁺ sensitive and high-gain system that ensures bursts of store-operated Ca²⁺ entry occur that are sufficient for the activation of *c-fos* expression, without the build-up of the Ca²⁺-dependent slow inactivation pathway that would abolish the response.

METHODS SUMMARY

Cultured RBL-1 cells and acutely isolated human nasal polyps were prepared as described previously⁷. Transfection with siRNA constructs was achieved using the AMAXA system. Patch-clamp recordings and cytoplasmic Ca²⁺ measurements followed standard methods⁷. RT-PCR, western blotting and confocal microscopy were carried out as described previously⁵. Error bars are s.e.m.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 April; accepted 23 November 2011.

Published online 9 January 2012.

- Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nature Rev. Mol. Cell Biol.* **3**, 639–650 (2002).
- Funk, C. D. Leukotriene modifiers as potential therapeutics for cardiovascular disease. *Nature Rev. Drug Discov.* **4**, 664–672 (2005).
- Naik, S. *et al.* Regulation of cysteinyl leukotriene type 1 receptor internalization and signaling. *J. Biol. Chem.* **280**, 8722–8732 (2005).
- Di Capite, J., Ng, S.-W. & Parekh, A. B. Decoding of cytoplasmic Ca²⁺ oscillations through the spatial signature drives gene expression. *Curr. Biol.* **19**, 853–858 (2009).
- Ng, S.-W., Nelson, C. & Parekh, A. B. Coupling of Ca²⁺ microdomains to spatially and temporally distinct cellular responses by the tyrosine kinase Syk. *J. Biol. Chem.* **284**, 24767–24772 (2009).
- Foletta, V. C., Segal, D. H. & Cohen, D. R. Transcriptional regulation in the immune system: all roads lead to AP-1. *J. Leukoc. Biol.* **63**, 139–152 (1998).
- Di Capite, J. L., Shirley, A., Nelson, C., Bates, G. & Parekh, A. B. Intercellular calcium wave propagation involving positive feedback between CRAC channels and cysteinyl leukotrienes. *FASEB J.* **23**, 894–905 (2009).
- Di Capite, J., Nelson, C., Bates, G. & Parekh, A. B. Targeting CRAC channels and leukotriene receptors provides a novel combination strategy for treating nasal polyposis. *J. Allergy Clin. Immunol.* **124**, 1014–1021 (2009).
- Stephan, V., Seibt, A., Koerholz, D. & Wahn, V. Expression of mRNA for the proto-oncogene *c-fos* in rat basophilic leukemia cells. *Cell. Signal.* **9**, 65–70 (1997).
- Jousset, H., Frieden, M. & Demaurex, N. STIM1 knockdown reveals that store-operated Ca²⁺ channels located close to sarco/endoplasmic Ca²⁺ ATPases (SERCA) pumps silently refill the endoplasmic reticulum. *J. Biol. Chem.* **282**, 11456–11464 (2007).
- Ozawa, K. *et al.* Ca²⁺-dependent and Ca²⁺-independent isozymes of protein kinase C mediate exocytosis in antigen-stimulated rat basophilic RBL-2H3 cells. Reconstitution of secretory responses with Ca²⁺ and purified isozymes in washed permeabilized cells. *J. Biol. Chem.* **268**, 1749–1756 (1993).
- Deshpande, D. A. *et al.* PKC-dependent regulation of the receptor locus dominates functional consequences of cysteinyl leukotriene type 1 receptor activation. *FASEB J.* **21**, 2335–2342 (2007).
- Bird, G. S. & Putney, J. W. J. Capacitative calcium entry supports calcium oscillations in human embryonic kidney cells. *J. Physiol.* **562**, 697–706 (2005).
- Stauffer, T. P., Ahn, S. & Meyer, T. Receptor-induced transient reduction in plasma membrane PtdIns(4,5)P₂ concentration monitored in living cells. *Curr. Biol.* **8**, 343–346 (1998).
- Parekh, A. B. & Putney, J. W. J. Store-operated calcium channels. *Physiol. Rev.* **85**, 757–810 (2005).
- Zweifach, A. & Lewis, R. S. Rapid inactivation of depletion-activated calcium current (ICRAC) due to local calcium feedback. *J. Gen. Physiol.* **105**, 209–226 (1995).

17. Fierro, L. & Parekh, A. B. Fast calcium-dependent inactivation of calcium release-activated calcium current (CRAC) in RBL-1 cells. *J. Membr. Biol.* **168**, 9–17 (1999).
18. Zweifach, A. & Lewis, R. S. Slow calcium-dependent inactivation of depletion-activated calcium current. *J. Biol. Chem.* **270**, 14445–14451 (1995).
19. Parekh, A. B. Slow feedback inhibition of calcium release-activated calcium current by calcium entry. *J. Biol. Chem.* **273**, 14925–14932 (1998).
20. Takasaka, T., Kaku, Y. & Hozawa, K. Mast cell degranulation in nasal polyps. *Acta Otolaryngol. Suppl.* **430**, 39–48 (1986).
21. Ozawa, K., Yamada, K., Kazanietz, M. G., Blumberg, P. M. & Beaven, M. A. Different isozymes of protein kinase C mediate feedback inhibition of phospholipase C and stimulatory signals for exocytosis in RBL-2H3 cells. *J. Biol. Chem.* **268**, 2280–2283 (1993).
22. Razin, E., Szallasi, Z., Kazanietz, M. G., Blumberg, P. M. & Rivera, J. Protein kinases C- β and C- ϵ link the mast cell high-affinity receptor for IgE to the expression of c-fos and c-jun. *Proc. Natl Acad. Sci. USA* **91**, 7722–7726 (1994).
23. Kasai, H. & Petersen, O. H. Spatial dynamics of second messengers: IP₃ and cAMP as long-range and associative messengers. *Trends Neurosci.* **17**, 95–101 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by Medical Research Council grant support to A.B.P. S.-W. N. held a BBSRC-Glaxo-Smith-Kline studentship. We thank G. Miesenboeck and J. Putney for comments on the manuscript.

Author Contributions S.-W.N. performed and analysed Ca²⁺ measurements, RT-PCR, western blots and siRNA knockdown. D.B. performed and analysed patch-clamp recordings. C.N. carried out immunocytochemistry. R.M. helped with Ca²⁺ measurements. R.A. and G.B. supplied human nasal polyps. A.B.P. carried out some patch-clamp experiments, discussed the results, supervised the project and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.B.P. (anant.parekh@dpag.ox.ac.uk).

METHODS

Cell culture and transfection. RBL-1 cells were bought from ATCC and were cultured (37 °C, 5% CO₂) in Dulbecco's modified Eagle medium with 10% fetal bovine serum, 2 mM L-glutamine and penicillin-streptomycin, as previously described²⁴. Cells were transfected with siRNA against Syk (purchased from Invitrogen; 5'-3' sense, CCCUCUGGCAGCUAGUGGAACAUA; antisense, UAAUGUCCACUAGCUGCCAGAGGG) using the Amaxa nucleofection system²⁵. siRNA against PKC isozymes were purchased from Invitrogen; PKC- α sense, GCGACACCUGCGACAUGAAUGUUA; antisense, UGAACAUAUCAU GUCGCAGGUGUCGC; PKC- β sense, GGCUAUGACCAACACCCAGGC AA; antisense, UUGCCUGGGUGUUGGUCAUUAGCC. Cells were used 36–48 h after plating.

Human tissue. Human polyps were removed from patients with full ethical consent and with approval from the National Research Ethics Service (REC number 07/H0607/104). Polyps were treated as described⁷ and used within 6 h of surgical removal.

ICRAC recordings. Patch-clamp experiments were conducted in the tight-seal whole-cell configuration at room temperature (20–24 °C) as previously described^{7,24}. Sylgard-coated, fire-polished pipettes had resistances of 4.2–5.5 M Ω when filled with standard internal solution that contained (in mM): Cs⁺ glutamate 145, NaCl 8, MgCl₂ 1, Mg-ATP 2, EGTA 10, CaCl₂ 4.6 mM (free Ca²⁺ ~140 nM), HEPES 10, pH 7.2 with CsOH. In some experiments, weak Ca²⁺ buffer was used instead (0.2 mM EGTA and no added CaCl₂). A correction of +10 mV was applied for the subsequent liquid junction potential that arose from this glutamate-based internal solution. The composition of the extracellular solution was (in mM): NaCl 145, KCl 2.8, CaCl₂ 10, MgCl₂ 2, CsCl 10, D-glucose 10, HEPES 10, pH 7.4 with NaOH.

I_{CRAC} was measured by applying voltage ramps (–100 to +100 mV in 50 ms) at 0.5 Hz from a holding potential of 0 mV. For fast inactivation, step pulses (250 ms duration) were applied from 0 mV to –100 mV every 2 s. Currents were filtered using an 8-pole Bessel filter at 2.5 kHz and digitized at 100 μ s. Currents were normalized by dividing the amplitudes (measured from the voltage ramps at –80 mV) by the cell capacitance. Capacitative currents were compensated before each ramp by using the automatic compensation of the EPC-9/2 (HEKA Elektronik) patch-clamp amplifier. For I_{CRAC} , leak currents were subtracted by averaging 2–3 ramp currents obtained just before I_{CRAC} had started to develop, and then subtracting this from all subsequent currents.

Ca²⁺ imaging. Ca²⁺ imaging experiments were carried out at room temperature (21–24 °C) using the IMAGO CCD camera-based system from TILL Photonics, as described previously²⁵. Cells were alternately excited at 356 and 380 nm (20 ms exposures) and images were acquired every 2 s. Images were analysed offline using IGOR Pro for Windows. Cells were loaded with Fura 2-AM (2 μ M) for 40 min at room temperature in the dark and then washed three times in standard external solution of composition (in mM) NaCl 145, KCl 2.8, CaCl₂ 2, MgCl₂ 2, D-glucose 10, HEPES 10, pH 7.4 with NaOH. Cells were left for 15 min to allow further de-esterification. Ca²⁺-free solution had the following composition (in mM) NaCl 145, KCl 2.8, MgCl₂ 2, D-glucose 10, HEPES 10, EGTA 0.1, pH 7.4 with NaOH).

EGTA-AM loading. Cells were loaded with EGTA by incubation for 45 min with EGTA-AM as described⁵.

Confocal microscopy. Cells were fixed in 4% paraformaldehyde in phosphate buffer for 30 min at room temperature. All the washes used 0.01% PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, 1 mM KH₂PO₄). The cells were blocked with 2% BSA and 10% goat serum for 1 h. Anti-c-Fos, -PKC- δ , -PKC- ϵ and -PKC- ζ were used in carrier (0.2% BSA, 1% goat serum) and left overnight at 4 °C and were purchased from Cell Signalling. Anti-PKC- α and - β were used in carrier (0.2% BSA, 1% goat serum) and left overnight at 4 °C and were purchased from Santa Cruz. The secondary anti-rabbit IgG was a HandL chain-specific (goat) fluorescein conjugate (Alexa Fluor 568, excitation at 578 nm, emission at 603 nm wavelength) from Invitrogen. This was used in PBS for 2 h at room temperature. The cells were mounted in Vectashield mounting medium. Images were obtained using a Leica confocal microscope, as described⁷.

Western blotting. Total cell lysates (50 μ g) were separated by SDS-PAGE on a 10% gel and electrophoretically transferred to nitrocellulose membrane, as described⁷. Membranes were blocked with 5% non-fat dry milk in TBS plus 0.1% Tween 20 (TBST) buffer for 1 h at room temperature. Membranes were washed with TBST three times and then incubated with primary antibody overnight at 4 °C. Anti-PKC- δ , - ϵ and - ζ antibodies were obtained from New England Biolabs and used at 1:2,500 dilution. Anti-PKC- α , - β and total ERK2 antibodies were purchased from Santa Cruz Biotechnology. PKC- α and - β antibodies were used at 1:2,500 dilution and total ERK antibody was used at a dilution of 1:5,000. The membranes were then washed with TBST again and incubated with a 1:2,500 dilution of goat anti-rabbit secondary antibody IgG from Santa Cruz Biotechnology for 1 h at room temperature. After washing with TBST, the bands were developed for visualization using ECL-plus western blotting detection system (GE Healthcare). Gels were quantified using the UN-SCAN-IT software package (Silk Scientific). Total ERK2 is widely used as a control for gel loading. The antibody does not discriminate between phosphorylated (and hence active) and non-phosphorylated ERK2 and therefore detects the total amount of this protein, regardless of whether the kinase has been activated. The extent of PKC was therefore normalized to the total amount of ERK2 present in each lysate, to correct for any differences in amount of cells used for each experiment.

RT-PCR. Total RNA was extracted from RBL cells by using an RNeasy Mini Kit (Qiagen), as described^{4,5}. RNA was quantified spectrophotometrically by absorbance at 260 nm. Total RNA (1 μ g) was reverse-transcribed using the iScriptTM cDNA Synthesis Kit (Bio-Rad), according to the manufacturer's instructions. Following cDNA synthesis, PCR amplification was then performed using BIO-X-ACTTM Short DNA Polymerase (Bioline) with primers specific for the detection of *c-fos* were synthesized by Invitrogen. The PCR products were electrophoresed through an agarose gel and visualized by ethidium bromide staining.

Statistics. Results are presented as means \pm s.e.m. Statistical significance was assessed using Student's *t*-test for comparison between two groups or ANOVA followed by a post-hoc Newman Keuls multiple comparison test for the difference between groups and results were considered significant at **P* < 0.05, ***P* < 0.01.

24. Glitsch, M. D., Bakowski, D. & Parekh, A. B. Activation of the store-operated Ca²⁺ current ICRAC is compromised by inhibitors of the lipoxygenase family of enzymes. *J. Physiol. (Lond.)* **539**, 93–106 (2002).
25. Ng, S.-W., Di Capite, J. L., Singaravelu, K. & Parekh, A. B. Sustained activation of the tyrosine kinase Syk by antigen in mast cells requires local Ca²⁺ influx through Ca²⁺ release-activated Ca²⁺ channels. *J. Biol. Chem.* **283**, 31348–31355 (2008).

Structures of cytochrome P450 17A1 with prostate cancer drugs abiraterone and TOK-001

Natasha M. DeVore¹ & Emily E. Scott¹

Cytochrome P450 17A1 (also known as CYP17A1 and cytochrome P450c17) catalyses the biosynthesis of androgens in humans¹. As prostate cancer cells proliferate in response to androgen steroids^{2,3}, CYP17A1 inhibition is a new strategy to prevent androgen synthesis and treat lethal metastatic castration-resistant prostate cancer⁴, but drug development has been hampered by lack of information regarding the structure of CYP17A1. Here we report X-ray crystal structures of CYP17A1, which were obtained in the presence of either abiraterone, a first-in-class steroidal inhibitor recently approved by the US Food and Drug Administration for late-stage prostate cancer⁵, or TOK-001, an inhibitor that is currently undergoing clinical trials^{4,6}. Both of these inhibitors bind the haem iron, forming a 60° angle above the haem plane and packing against the central I helix with the 3β-OH interacting with asparagine 202 in the F helix. Notably, this binding mode differs substantially from those that are predicted by homology models and from steroids in other cytochrome P450 enzymes with known structures, and some features of this binding mode are more similar to steroid receptors. Whereas the overall structure of CYP17A1 provides a rationale for understanding many mutations that are found in patients with steroidogenic diseases, the active site reveals multiple steric and hydrogen bonding features that will facilitate a better understanding of the enzyme's dual hydroxylase and lyase catalytic capabilities and assist in rational drug design. Specifically, structure-based design is expected to aid development of inhibitors that bind only CYP17A1 and solely inhibit its androgen-generating lyase activity to improve treatment of prostate and other hormone-responsive cancers.

CYP17A1 is a membrane-bound dual-function monooxygenase with a critical role in the synthesis of many human steroid hormones¹. The 17α-hydroxylase activity of CYP17A1 is required for the generation of glucocorticoids such as cortisol, but both the hydroxylase and 17,20-lyase activities of CYP17A1 are required for the production of androgenic and oestrogenic sex steroids (Supplementary Fig. 1). CYP17A1 is an important target for the treatment of breast and prostate cancers that proliferate in response to oestrogens and androgens^{2,3}. In the absence of structural information, CYP17A1 inhibitors have been

designed that are thought to bind the cytochrome P450 haem iron⁴, but it has been difficult to rationalize or predict other structural features that are critical for effective and selective CYP17A1 inhibition. In addition, structural information is important for understanding 17-hydroxylase deficiencies and may also improve our understanding of polycystic ovary disease⁷. We determined the structures of human CYP17A1 when bound to two clinically relevant CYP17A1 inhibitors (Supplementary Fig. 2). Abiraterone is the active form of a prodrug that was recently approved by the US Food and Drug Administration for metastatic prostate cancer^{5,8} and is under investigation for breast cancer⁹. TOK-001 is currently undergoing clinical trials for prostate cancer⁴.

A truncated, His-tagged version of the human CYP17A1 protein was generated from a synthetic complementary DNA that was engineered to remove the single amino-terminal transmembrane helix, and this CYP17A1 was expressed in *Escherichia coli*. The resulting CYP17A1 was membrane bound and so it was solubilized with detergent before purification. This CYP17A1 binds abiraterone (17-(3-pyridyl)androst-5,16-dien-3β-ol) (Fig. 1a) and TOK-001 (17-(1H-Benzimidazol-1-yl)androst-5,16-dien-3β-ol) with decreases in absorbance at 402 nm and increases in absorbance at 424 nm, consistent with nitrogen binding to the haem iron (type II interaction) with dissociation constant (K_d) values of <100 nM (Fig. 1a, inset). Similar titrations with the substrates progesterone (Fig. 1b) and pregnenolone revealed decreases in absorbance at 419 nm and increases in absorbance at 385 nm, which is indicative of the ligand displacing water from the haem (type I interaction). CYP17A1 binds pregnenolone (K_d <100 nM) more tightly than progesterone (K_d 229 ± 14 nM; Fig. 1b, inset). We found that full-length enzyme¹⁰ had a similar k_{cat} and three-fold higher Michaelis constant (k_{cat} 1.31 ± 0.03 min⁻¹, K_m 11.4 ± 0.7 μM) compared to the truncated form (k_{cat} 1.31 ± 0.03 min⁻¹, K_m 3.7 ± 0.3 μM). The half-maximum inhibitory concentration (IC₅₀) values for abiraterone (201 ± 1 nM) were lower than for TOK-001 (503 ± 1 nM) (Fig. 1c). Thus, truncated human CYP17A1 is a functional enzyme in terms of ligand binding, catalytic function and inhibition.

Both structures with abiraterone (2.6 Å) and TOK-001 (2.4 Å) demonstrate the characteristic cytochrome P450 fold (Fig. 2a) and have four very similar protein copies in each asymmetric unit

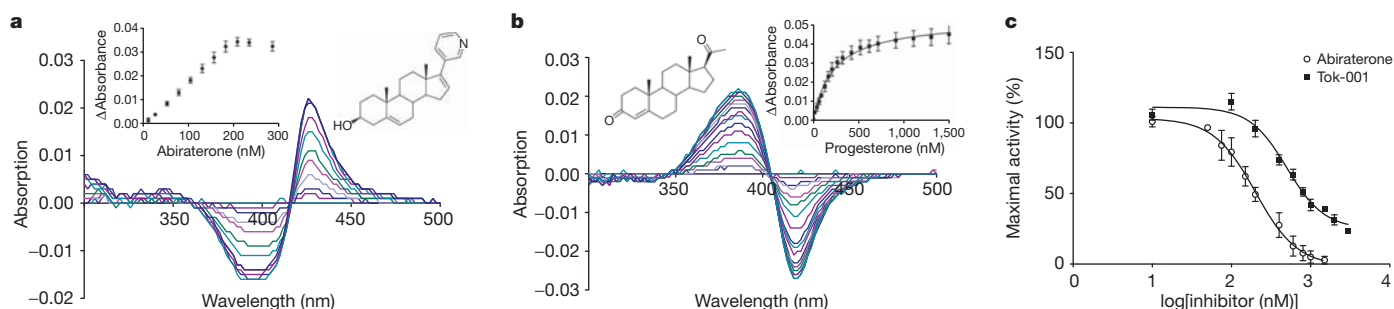


Figure 1 | Function of CYP17A1 and inhibition by clinical compounds. **a**, CYP17A1 titration with abiraterone (10–274 nM) yields progressive shifts in the ultraviolet–visible difference spectrum that indicate nitrogen binding to

haem iron. **b**, Similar titration with progesterone (10–1535 nM) indicates water displacement from the haem iron. **c**, IC₅₀ of abiraterone (circles) and TOK-001 (squares) for progesterone 17α-hydroxylation.

¹Department of Medicinal Chemistry, 1251 Wescoe Hall Drive, University of Kansas, Lawrence, Kansas 66045, USA.

(Supplementary Table 1). Consistent with spectral binding data, abiraterone and TOK-001 bind with the nitrogen of the C17 pyridine or benzimidazole, respectively, forming a coordinate covalent bond with the haem iron (Fig. 2b, d). The steroidal cores of these inhibitors rise at an angle of 60° above the haem plane, directed between the F and G helices (Fig. 2b, d), and essentially overlap with each other (Fig. 2f). The unsubstituted α -face packs flat against the I helix where G301, A302 and adjacent residues form a highly complementary hydrophobic planar surface (Fig. 2b). The 3β -OH groups of abiraterone (Fig. 2b) and TOK-001 (Fig. 2d) hydrogen bond with N202 in the F helix (~ 2.6 Å and ~ 2.4 Å, respectively).

Although inhibitors occupy the majority of the enclosed active site, the void extends beyond these ligands in several directions. First, the active site wall nearest the inhibitor β -face is not as complementary to the steroid core as to the α -face. The C18 and C19 methyl groups project towards a crevice between the B' helix, the $\beta 4$ loop and the loop following the F helix (Fig. 2b). Only three side chains of the cavity wall come within 4 Å of C18 or C19. The cavity wall facing the β -face of abiraterone or TOK-001 is primarily lined with hydrophobic atoms of A105, S106, A113, F114, I206, L209, V236 and V482 (Fig. 2c), but there are two notable exceptions. R239 and D298 extend from the G and I helices, respectively, to orient their basic and acidic termini towards C6. These two polar side chains flank a substantial extension of the active site void adjacent to C6. Second, in the abiraterone structure there is unfilled space in the active site cavity adjacent to the pyridine

ring bordered by V366, A367, I371 and V483 (Fig. 2c), and this volume is occupied by benzimidazole in the TOK-001 structure (Fig. 2e). Last, the most substantial active site cavity extension is from the 3β -OH of the inhibitors over the top of helix I and along the underside of helices F and G. This cavity is mostly lined by hydrophobic residues (I198, L243 and F300), but its 'roof' is bordered by several polar F and G helix residues (Y201, N202 and R239; Fig. 2c) that interact with, or are located near, water molecules in this region. The overall cavity containing TOK-001 is similar to that observed when abiraterone is bound, except that the available void volume over helix I is slightly smaller (Fig. 2e).

The single direct hydrogen bond between inhibitors and the protein is part of a larger hydrogen bonding network. In the abiraterone complex this network involves N202, E305, several conserved water molecules, R239, the backbone carbonyl of G297 and, in some molecules, Y201 (Fig. 3). Although Y201 is not within hydrogen bonding distance to these waters for molecules A and B, the side chain rotates slightly towards abiraterone in molecules C and D to interact with one or both of the water molecules. TOK-001 has a very similar hydrogen bonding network (Supplementary Fig. 3). These interactions are strongly reminiscent of the interactions that are conserved in the androgen, oestrogen, glucocorticoid, mineralocorticoid and progesterone receptors¹¹ (Fig. 4a, b). In each receptor, the 3β -OH or 3-keto group of steroids binds within a deep receptor pocket and forms hydrogen bonds with a glutamine or glutamate, an arginine and often a conserved water

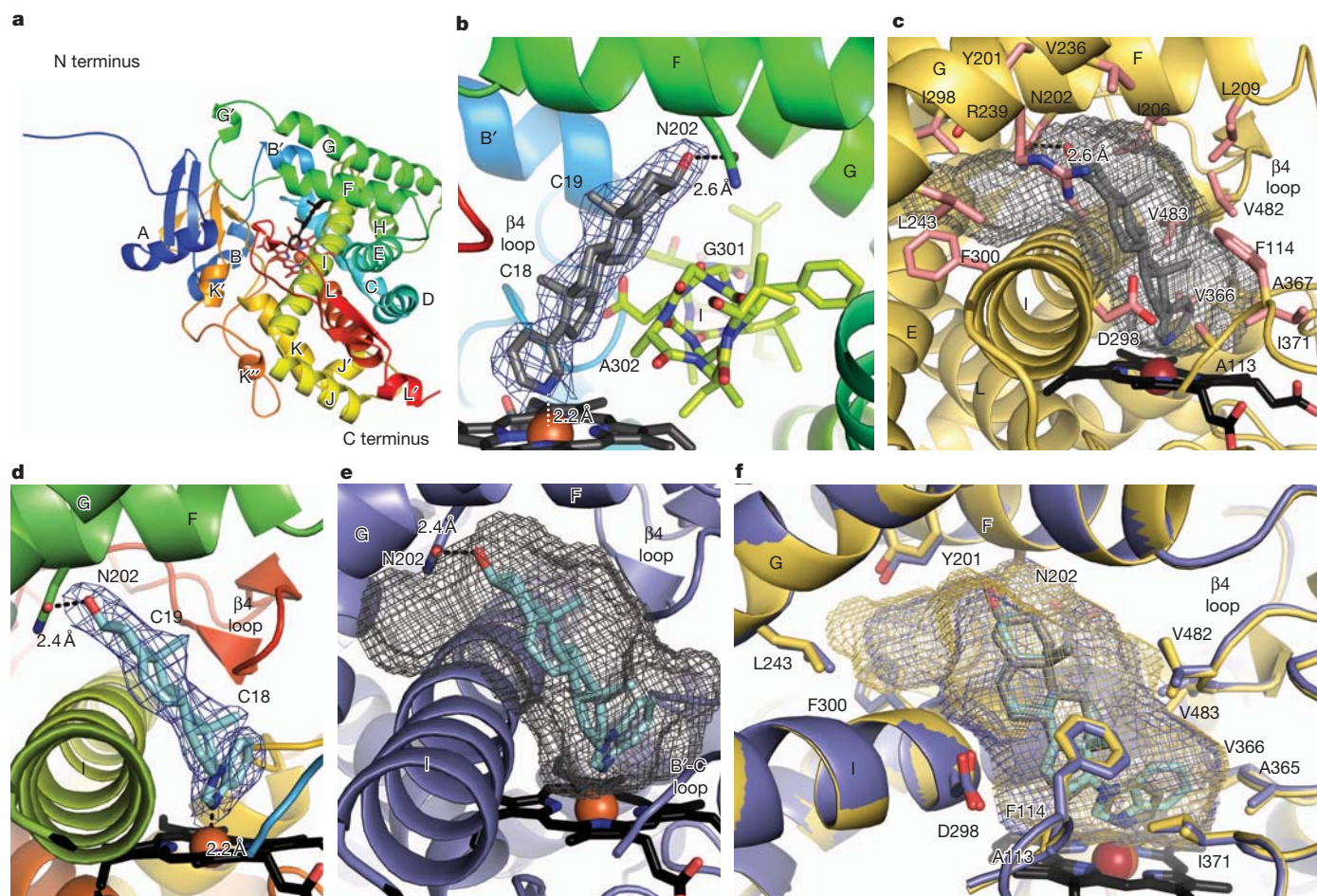


Figure 2 | CYP17A1 ligand binding. In the stick and sphere representations, non-carbon atoms are indicated in blue (N), red (O), black (haem), grey (abiraterone), cyan (TOK-001) and by a dark red sphere (Fe). **a**, A coloured representation of the CYP17A1–abiraterone structure, from the N terminus (blue) to the C terminus (red). **b**, Abiraterone binds ($2|F_o| - |F_c|$ density at 1σ ; blue mesh) at an angle of approximately 60° from haem against helix I (yellow).

c, Abiraterone cavity (grey mesh), rotated approximately 180° in relation to **b**. **d**, TOK-001 binding ($2|F_o| - |F_c|$ density at 1σ , blue mesh). **e**, TOK-001 cavity (grey mesh). **f**, Abiraterone structure (yellow) superimposed on TOK-001 structure (purple) with respective voids (mesh). The B' helix has been removed from panels **c–f** so that it is possible to view the ligands.

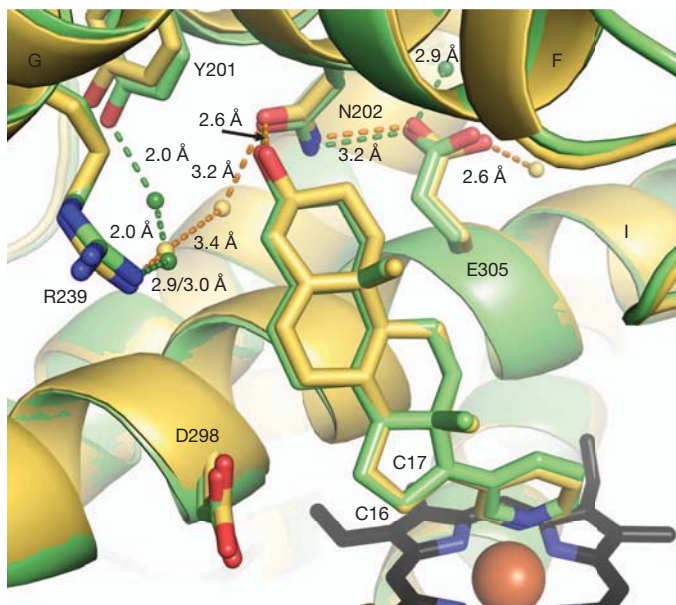


Figure 3 | Hydrogen bond network with abiraterone. CYP17A1 has a hydrogen bonding network at the top of the active site that interacts with abiraterone. Molecules A and B (yellow), and C and D (green) have slightly different networks, with the main difference being the involvement of Y201. Water molecules are indicated by small spheres. Hydrogen bonds are represented by dashed lines and the distances between them are indicated.

molecule. These interactions are critical for ligand recognition by hormone receptors¹² and may also contribute to CYP17A1 selectivity for pregnenolone, progesterone and their 17 α -hydroxy derivatives. Notably, TOK-001 is both a CYP17A1 inhibitor and androgen receptor

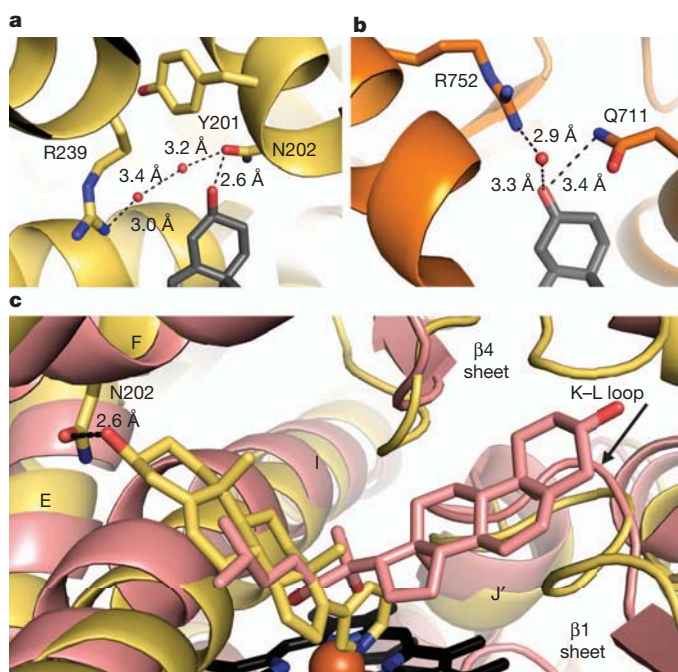


Figure 4 | CYP17A1 compared to the androgen receptor and CYP11A1. **a**, The hydrogen bonding network near the abiraterone 3 β -OH involves N202, R239 and conserved waters. **b**, The androgen receptor (PDB 3L3X) has a similar hydrogen bond network with R752, Q711 and several waters that interact with the dihydrotestosterone ketone. **c**, The structure of CYP17A1 with abiraterone (yellow) superimposed on the structure of CYP11A1 with 20,22-dihydroxycholesterol (PDB 3NA0, pink). This shows that there are markedly different steroid orientations between the two structures.

antagonist¹³, and the similarity of these binding modes is probably the reason for this dual mechanism of action.

Orientations of the native CYP17A1 substrates are of substantial value in understanding the function of this enzyme. Pregnenolone and progesterone were docked into the CYP17A1–abiraterone structure modelled as the Fe(IV)=O (compound I) catalytic state. Progesterone maintained the N202 hydrogen bond. The distances from C17 and C16 to the catalytic oxygen were 3.7 Å and 3.9 Å, respectively, distances that are consistent with the experimentally observed major 17 α -OH and minor 16 α -OH progesterone metabolites. The pregnenolone 3 β -OH hydrogen was also bonded to N202 and the C17 atom was 3.6 Å from the Compound I oxygen (Supplementary Fig. 4). However, the active site topology may be altered in the presence of substrates and this is an important area for further investigation.

CYP17A1 can be compared to three other P450 enzymes that are involved in steroidogenesis or cholesterol metabolism with reported steroid complex structures: CYP19A1 (aromatase)¹⁴, CYP11A1 (cholesterol side-chain cleavage enzyme)¹⁵ and CYP46A1, a cholesterol 24-hydroxylase¹⁶. Although all four enzymes maintain the canonical cytochrome P450 fold, CYP19A1, CYP11A1 and CYP46A1 orient steroids in the opposite direction from CYP17A1 so that the steroid ligands are positioned over the K–L loop directed towards the β 1 sheet (as shown for CYP11A1 in Fig. 4c) instead of oriented towards helices F and G.

More than 50 CYP17A1 mutations have been identified, and most of these were identified in patients with 17-hydroxylase deficiencies. The biochemical effects of many clinical mutations can be rationalized by examining the CYP17A1 structure (Supplementary Fig. 5 and Supplementary Table 2). Mutations R96W, R125Q, H373D/N and R440H/C all alter residues that directly interact with the haem propionates and probably disrupt haem binding, resulting in the complete loss of activity that was observed^{17–22}. Mutations E305G, R347H/C, R358Q and R449A eliminate only the lyase activity of CYP17A1^{23–25}. E305 forms a hydrogen bond to N202 in the active site, indicating a role in substrate positioning, whereas the other residues are on the proximal face of the protein (Supplementary Fig. 5), and this is consistent with their proposed role in cytochrome *b*₅ binding^{23,26}, which promotes lyase activity. Finally, reduction of the minor 16 α -hydroxyprogesterone metabolite is reported for the artificial mutation A105L²⁷, consistent with its location in the active site facing the β -face where the additional bulk may reduce the steroid movement within the active site.

Abiraterone and TOK-001 have several features that make them effective inhibitors of CYP17A1: first, a heterocyclic nitrogen that coordinates to the haem iron; second, a planar α -face to pack against the I helix; and third, hydrogen bonding interactions of 3 β -OH with conserved polar residues in a hydrogen binding network. These structures provide a model for the binding of substrates and other inhibitors that is very different from binding orientations that have previously been proposed by homology modelling and docking studies and from structures that have been demonstrated in other steroid-metabolizing cytochrome P450 enzymes. Perhaps most importantly, the cavity that is observed is not bilobed as predicted by many modelling studies^{28,29}. CYP17A1 interactions with these inhibitors are instead more reminiscent of steroid binding to steroid receptors, and this may be the origin of the TOK-001 dual mechanism of action. Thus, these structures contribute to a better understanding of the function and inhibition of CYP17A1 in a way that should substantially benefit the apprehension of enzyme dysfunction in clinical disease and enable structure-based drug design of CYP17A1 inhibitors for treating hormone-responsive cancers, particularly prostate cancer.

METHODS SUMMARY

A synthetic cDNA for human CYP17A1 was modified to delete residues 2–19, substitute the hydrophilic sequence Arg 20–Arg 21–Cys 22–Pro 23 with Ala 20–Lys 21–Lys 22–Thr 23 and add a carboxy-terminal four-histidine tag (Supplementary Fig. 6) before cloning into the pCWori⁺ plasmid and overexpression in *Escherichia*

coli JM109 cells. Protein was purified by nickel affinity, cation exchange and size exclusion chromatography. Abiraterone was synthesized (see Methods). Binding affinities were determined using an ultraviolet–visible spectral shift assay. Progesterone 17 α -hydroxylation was evaluated using HPLC separation and ultraviolet detection. For crystallography, inhibitors were included throughout purification. Crystals were grown from CYP17A1 (30 mg ml⁻¹) complexed with inhibitor using hanging-drop vapour diffusion to equilibrate against 30% PEG 3350, 0.175 M Tris, pH 8.5, 0.30 M ammonium sulphate and 3% glycerol. Diffraction data were collected and phased by molecular replacement. Iterative model building and refinement generated the final model. Substrates were docked using Surflex-Dock.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 May; accepted 30 November 2011.

Published online 22 January 2012.

- Miller, W. L. & Auchus, R. J. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr. Rev.* **32**, 81–151 (2011).
- Attard, G., Reid, A. H., Olmos, D. & de Bono, J. S. Antitumor activity with CYP17 blockade indicates that castration-resistant prostate cancer frequently remains hormone driven. *Cancer Res.* **69**, 4937–4940 (2009).
- Yap, T. A., Carden, C. P., Attard, G. & de Bono, J. S. Targeting CYP17: Established and novel approaches in prostate cancer. *Curr. Opin. Pharmacol.* **8**, 449–457 (2008).
- Vasaitis, T. S., Bruno, R. D. & Njar, V. C. CYP17 inhibitors for prostate cancer therapy. *J. Steroid Biochem. Mol. Biol.* **125**, 23–31 (2011).
- de Bono, J. S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* **364**, 1995–2005 (2011).
- Molina, A. & Belidegrun, A. Novel therapeutic strategies for castration resistant prostate cancer: inhibition of persistent androgen production and androgen receptor mediated signaling. *J. Urol.* **185**, 787–794 (2011).
- Auchus, R. J., Geller, D. H., Lee, T. C. & Miller, W. L. The regulation of human P450c17 activity: relationship to premature adrenarche, insulin resistance and the polycystic ovary syndrome. *Trends Endocrinol. Metab.* **9**, 47–50 (1998).
- Attard, G. *et al.* Phase I clinical trial of a selective inhibitor of CYP17, abiraterone acetate, confirms that castration-resistant prostate cancer commonly remains hormone driven. *J. Clin. Oncol.* **26**, 4563–4571 (2008).
- Brodie, A., Njar, V., Macedo, L. F., Vasitis, T. S. & Sabnis, G. The Coffey Lecture: steroidogenic enzyme inhibitors and hormone dependent cancer. *Urol. Oncol.* **27**, 53–63 (2009).
- Imai, T. *et al.* Expression and purification of functional human 17 α -hydroxylase/17,20-lyase (P450c17) in *Escherichia coli*. Use of this system for study of a novel form of combined 17 α -hydroxylase/17,20-lyase deficiency. *J. Biol. Chem.* **268**, 19681–19689 (1993).
- Huang, P., Chandra, V. & Rastinejad, F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu. Rev. Physiol.* **72**, 247–272 (2010).
- Pereira de Jesus-Tran, K. *et al.* Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. *Protein Sci.* **15**, 987–999 (2006).
- Vasaitis, T. *et al.* Androgen receptor inactivation contributes to antitumor efficacy of 17 α -hydroxylase/17,20-lyase inhibitor 3 β -hydroxy-17-(1H-benzimidazole-1-yl)androsta-5,16-diene in prostate cancer. *Mol. Cancer Ther.* **7**, 2348–2357 (2008).
- Ghosh, D., Griswold, J., Erman, M. & Pangborn, W. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature* **457**, 219–223 (2009).
- Mast, N. *et al.* Structural basis for three-step sequential catalysis by the cholesterol side chain cleavage enzyme CYP11A1. *J. Biol. Chem.* **286**, 5607–5613 (2011).
- Mast, N. *et al.* Crystal structures of substrate-bound and substrate-free cytochrome P450 46A1, the principal cholesterol hydroxylase in the brain. *Proc. Natl Acad. Sci. USA* **105**, 9546–9551 (2008).
- Dhir, V. *et al.* Steroid 17 α -hydroxylase deficiency: Functional characterization of four mutations (A174E, V178D, R440C, L465P) in the CYP17A1 gene. *J. Clin. Endocrinol. Metab.* **94**, 3058–3064 (2009).
- Rosa, S. *et al.* Clinical, genetic and functional characteristics of three novel CYP17A1 mutations causing combined 17 α -hydroxylase/17,20-lyase deficiency. *Horm. Res. Paediatr.* **73**, 198–204 (2010).
- Katsumata, N., Ogawa, E., Fujiwara, I. & Fujikura, K. Novel CYP17A1 mutation in a Japanese patient with combined 17 α -hydroxylase/17,20-lyase deficiency. *Metabolism* **59**, 275–278 (2010).
- Ergun-Longmire, B. *et al.* Two novel mutations found in a patient with 17 α -hydroxylase enzyme deficiency. *J. Clin. Endocrinol. Metab.* **91**, 4179–4182 (2006).
- Sahakitrungruang, T., Tee, M. K., Speiser, P. W. & Miller, W. L. Novel P450c17 mutation H373D causing combined 17 α -hydroxylase/17,20-lyase deficiency. *J. Clin. Endocrinol. Metab.* **94**, 3089–3092 (2009).
- Biason-Lauber, A. *et al.* 17 α -hydroxylase/17,20-lyase deficiency as a model to study enzymatic activity regulation: role of phosphorylation. *J. Clin. Endocrinol. Metab.* **85**, 1226–1231 (2000).
- Lee-Robichaud, P. *et al.* The cationic charges on Arg347, Arg358 and Arg449 of human cytochrome P450c17 (CYP17) are essential for the enzyme's cytochrome b₅-dependent acyl-carbon cleavage activities. *J. Steroid Biochem. Mol. Biol.* **92**, 119–130 (2004).
- Gupta, M. K., Geller, D. H. & Auchus, R. J. Pitfalls in characterizing P450c17 mutations associated with isolated 17,20-lyase deficiency. *J. Clin. Endocrinol. Metab.* **86**, 4416–4423 (2001).
- Tiosano, D. *et al.* Metabolic evidence for impaired 17 α -hydroxylase activity in a kindred bearing the E305G mutation for isolate 17,20-lyase activity. *Eur. J. Endocrinol.* **158**, 385–392 (2008).
- Auchus, R. J., Lee, T. C. & Miller, W. L. Cytochrome b₅ augments the 17,20-lyase activity of human P450c17 without direct electron transfer. *J. Biol. Chem.* **273**, 3158–3165 (1998).
- Swart, A. C., Storbeck, K. H. & Swart, P. A single amino acid residue, Ala 105, confers 16 α -hydroxylase activity to human cytochrome P450 17 α -hydroxylase/17,20-lyase. *J. Steroid Biochem. Mol. Biol.* **119**, 112–120 (2010).
- Haider, S. M., Patel, J. S., Poojari, C. S. & Neidle, S. Molecular modeling on inhibitor complexes and active-site dynamics of cytochrome P450 C17, a target for prostate cancer therapy. *J. Mol. Biol.* **400**, 1078–1098 (2010).
- Jagus, C. *et al.* Synthesis, biological evaluation and molecular modelling studies of methyleneimidazole substituted biaryls as inhibitors of human 17 α -hydroxylase-17,20-lyase (CYP17). Part I: heterocyclic modifications of the core structure. *Bioorg. Med. Chem.* **16**, 1992–2010 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements X-ray data were collected at the Stanford Synchrotron Radiation Lightsource (SSRL). The SSRL Structural Molecular Biology Program is supported by the US Department of Energy Office of Biological and Environmental Research and by the US National Institutes of Health (NIH), National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. We thank C.-J. Liu and the University of Kansas (KU) Center of Biomedical Research Excellence (COBRE) Center for Cancer Experimental Therapeutics for synthesizing abiraterone (NIH RR030926), M. R. Waterman for the full-length CYP17A1 construct, J. Wang for assistance with the Fe(IV)=O construct used in docking and A. Skinner and J. Aubé for manuscript suggestions. This research was funded by the NIH through the KU COBRE Center for Protein Structure and Function (NIH RR17708) and GM076343.

Author Contributions N.M.D. engineered, expressed, characterized, purified and crystallized CYP17A1 under the direction of E.E.S. N.M.D. and E.E.S. jointly performed X-ray diffraction experiments, solved and refined the structures, and wrote the manuscript. N.M.D. performed the docking studies of CYP17A1.

Author Information Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under the accession codes 3RUK for CYP17A1 with abiraterone and 3SWZ for CYP17A1 with TOK-001. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.E.S. (eescott@ku.edu).

METHODS

Synthesis and characterization of abiraterone. A stirred solution of 17-iodoandrosta-5,16-dien-3 β -ol (600 mg, 1.5 mmol) in tetrahydrofuran (20 ml) in a 100-ml round-bottomed flask was purged with argon. Bis(triphenylphosphine) palladium (II) chloride catalyst (11 mg, 0.016 mmol) and then diethyl(3-pyridyl) borane (265 mg, 1.8 mmol) were added. To the resultant orange solution, an aqueous solution of sodium carbonate (2 M, 5 ml) was added. The flask was fitted with a reflux condenser and the apparatus was purged again with argon. The mixture was then heated under reflux (at $\sim 80^\circ\text{C}$) and stirred for 4 days then allowed to cool. The mixture was poured into water and extracted with hot toluene (3×30 ml). The toluene extracts were dried (using Na_2CO_3) and concentrated. Column chromatography was performed with Et_2O and toluene (with a ratio of 1:2) as the eluent to give abiraterone (350 mg, 66%) as a white crystalline solid: melting point $228\text{--}230^\circ\text{C}$; IR O-H stretching frequency ν_{max} 3307 cm^{-1} ; ^1H NMR peak shifts in ppm (δ) 1.07 (s, 3, H-19), 1.09 (s, 3, H-18), 3.54 (m, 1, H-3 α), 5.41 (dm, 1, J = 5.2 Hz, H-6), 6.01 (m, 1, H-16), 7.24 (dd, 1, pyridyl H-5), 7.66 (dd, 1, pyridyl H-4), 8.47 (dd, 1, pyridyl H-6), 8.63 (d, 1, pyridyl H-2); ^{13}C NMR δ 151.69, 147.92, 147.84, 141.19, 133.68, 132.98, 129.24, 123.03, 121.32, 71.65, 57.56, 50.36, 47.34, 42.32, 37.19, 36.71, 35.26, 31.81, 31.64, 30.45, 20.88, 19.35, 16.59. The high resolution mass spectrum calculated the mass to charge ratio (m/z) $\text{C}_{24}\text{H}_{32}\text{NO}$ $[\text{M} + \text{H}]^+$ to be 350.2484. The experimental value was 350.2491. Abiraterone was 99% pure by liquid chromatography mass spectrometry.

CYP17A1 design, expression and purification. The human CYP17A1 cDNA was synthesized with codon optimization for *E. coli* expression (Blue Heron Biotechnology). A truncated and His-tagged construct was generated by truncation of the N-terminal transmembrane helix ($\Delta 2\text{--}19$), substitution of Arg 20-Arg 21-Cys 22-Pro 23 with Ala 20-Lys 21-Lys 22-Thr 23, and addition of a C-terminal four-residue histidine tag (Supplementary Fig. 6). N-terminal modifications were designed to increase solubility. This altered cDNA was inserted into the pCWori⁺ expression vector and expressed in *E. coli* JM109 cells.

Cells were grown, collected and disrupted, as described previously^{10,30}. After centrifugation (5,000g), CYP17A1 was solubilized with either 4.8 mM Cymal-5 (for crystallography; Affymetrix) or 2% Emulgen 913 (for assays; Desert Biologicals) and then underwent ultracentrifugation (80,000g) for 60 min. The lysate was loaded onto a nickel nitrilotriacetic acid agarose (Qiagen) column and purified, as reported elsewhere³⁰. Eluted CYP17A1 fractions were pooled, diluted fivefold with buffer A (50 mM Tris, pH 7.4, 20% glycerol, 100 mM glycine and 1 mM EDTA) and loaded on a HiTrap CM fast flow column (GE Healthcare). Protein was eluted in buffer A with 0.5 M NaCl, concentrated to ~ 1 ml, and loaded on a Superdex 200 16/60 size exclusion column (GE Healthcare). For crystallography, 10 μM abiraterone or TOK-001 (Shanghai Haoyuan Chemexpress) was added to all buffers.

Protein crystallization, data collection and structure determination. CYP17A1 crystals were grown using hanging-drop vapour diffusion equilibration. CYP17A1 (30 mg ml^{-1}) in buffer A supplemented with 10 μM abiraterone or TOK-001, 2.4 mM Cymal-5 and 0.5 M NaCl was equilibrated against 30% PEG 3350, 0.175 M Tris, pH 8.5, 0.30 M ammonium sulphate and 3% glycerol at 20°C . Crystals were cryoprotected in 7:1 mother liquor:ethylene glycol and flash cooled in liquid nitrogen. Native data sets were collected at 0.98 Å, 100 K at the Stanford Synchrotron Radiation Laboratory beamline 9-2. Data were processed using Mosflm³¹ and Scala³². The abiraterone complex was solved by molecular replacement using BALBES³³ with a final search model based on CYP2R1 (PDB 3CZH), and the TOK-001 structure was solved using Phaser³² with the abiraterone structure as a search model. Model building and refinement were accomplished iteratively using COOT³⁴ and Refmac³⁵ in CCP4³². Structure validation was performed using WHATCHECK³⁶ and PROCHECK³⁷. Ramachandran plot analysis reveals per cent favourable/additional allowed/generously allowed/disallowed residues are 86.3/13.2/0.5/0.0 (abiraterone structure) and 86.2/13.3/0.5/0.0 (TOK-001 structure). X-ray statistics are provided (Supplementary Table 1). Probe-occupied voids were

calculated using VOIDOO³⁸ (probe radius = 1.4 Å; grid mesh = 0.4 Å). All figures were prepared using MacPyMOL³⁹.

Docking. The CYP17A1 active site was defined as described for other cytochromes P450 (ref. 40) with the addition of an oxygen molecule directly coordinated to the haem to mimic compound I of the cytochrome P450 catalytic cycle. Substrate coordinates were prepared and energy was minimized with SYBYL (Tripos International). Charges were assigned using the Gasteiger and Marsili method. Surflex-Dock⁴¹ (Tripos International) was used to dock ligands as previously described³⁸. The active site was a 10 Å sphere around the haem and pregnenolone. Movement of pregnenolone within the active site was not substantial, with the distances from C17 to O=Fe(IV) as 4.5 Å, from C16 to O=Fe(IV) as 4.5 Å, and from C21 to O=Fe(IV) as 3.0 Å for the lowest energy pose.

Enzyme activity and IC₅₀ determinations. Progesterone 17 α -hydroxylation was evaluated using a modified HPLC method with ultraviolet detection⁴². CYP17A1 (50 pmol) and rat NADPH-cytochrome P450 reductase⁴³ were mixed with a ratio of 1:4, incubated on ice (20 min) and added to buffer (50 mM Tris, pH 7.4 and 5 mM MgCl_2) containing progesterone (0–50 μM) to a total volume of 500 μl . Phosphatidylcholine (25 μg) was included for side-by-side kinetic comparisons with the full-length enzyme¹⁰. For IC₅₀ determinations, inhibitor concentrations were 0–1.5 μM for abiraterone and 0–3.0 μM for TOK-001. After warming (37°C , 3 min), reactions were initiated by the addition of NADPH (20 μl 25 mM), incubated for 10 min (37°C), quenched with 20% trichloroacetic acid (300 μl) and placed on ice. The 17 α -hydroxyprogesterone metabolite was identified by ultraviolet detection at 248 nm after HPLC separation and co-eluted with authentic standards. The HPLC mobile phase was 40% acetonitrile, 60% water with 1% acetic acid and run at 1 ml min^{-1} (Luna (Phenomenex), 5 μm , C18, 50×4.6 mm).

Ligand binding assays. Ligand binding assays were based on spectral differences that were detected after ligand titration and were performed as described³⁸, except that the CYP17A1 concentration was 0.1 μM , the path length was 5 cm and the tight binding equation was used.

Functional data were analysed using Prism (GraphPad Software) and presented as mean \pm standard error.

- Pechurskaya, T. A., Lukashovich, O. P., Gilep, A. A. & Usanov, S. A. Engineering, expression, and purification of “soluble” human cytochrome P45017 α and its functional characterization. *Biochemistry* **73**, 806–811 (2008).
- Leslie, A. G. W. MOSFLM 6.0 (Cambridge, 1998).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. BALBES: a molecular replacement pipeline. *Acta Crystallogr. D* **64**, 125–132 (2008).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
- Hoof, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
- DeVore, N. M. et al. Key residues controlling binding of diverse ligands to human cytochrome P450 2A enzymes. *Drug Metab. Dispos.* **37**, 1319–1327 (2009).
- DeLano, W. L. The PyMol Molecular Graphics System (DeLano Scientific, 2002).
- Kleywegt, G. J. & Jones, T. A. Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr. D* **50**, 178–185 (1994).
- Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **46**, 499–511 (2003).
- Hutschenreuter, T. U., Ehmer, P. B. & Hartmann, R. W. Synthesis of hydroxy derivatives of highly potent non-steroidal CYP 17 inhibitors as potential metabolites and evaluation of their activity by a non cellular assay using recombinant human enzyme. *J. Enzyme Inhib. Med. Chem.* **19**, 17–32 (2004).
- Shen, A. L., Porter, T. D., Wilson, T. E. & Kasper, C. B. Structural analysis of the FMN binding domain of NADPH-cytochrome P-450 oxidoreductase by site-directed mutagenesis. *J. Biol. Chem.* **264**, 7584–7589 (1989).

CORRECTIONS & AMENDMENTS

CORRIGENDUM

doi:10.1038/nature10788

A reserve stem cell population in small intestine renders *Lgr5*-positive cells dispensable

Hua Tian, Brian Biehs, Søren Warming, Kevin G. Leong, Linda Rangell, Ophir D. Klein & Frederic J. de Sauvage

Nature **478**, 255–259 (2011).

In this Letter, two additional references^{32,33}, which described stem-cell organization in the intestinal crypt, should have been cited after the first sentence of the main text: “Two types of stem cells have been described in the small intestine based on location and cycling dynamics^{1–4,32,33}”. The references and their citations have been added to the PDF and HTML versions.

32. Potten, C. S., Gandara, R., Mahida, Y. R., Loeffler, M. & Wright, N. A. The stem cells of small intestinal crypts: where are they? *Cell Prolif.* **42**, 731–750 (2009).
33. Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. I. Evidence from Paneth cells in the adult mouse. *Am. J. Anat.* **160**, 51–63 (1981).

CORRECTIONS & AMENDMENTS

CORRIGENDUM

doi:10.1038/nature10817

The unusual γ -ray burst GRB 101225A from a helium star/ neutron star merger at redshift 0.33

C. C. Thöne, A. de Ugarte Postigo, C. L. Fryer, K. L. Page,
J. Gorosabel, M. A. Aloy, D. A. Perley, C. Kouveliotou,
H. T. Janka, P. Mimica, J. L. Racusin, H. Krimm, J. Cummings,
S. R. Oates, S. T. Holland, M. H. Siegel, M. De Pasquale,
E. Sonbas, M. Im, W.-K. Park, D. A. Kann, S. Guziy,
L. Hernández García, A. Llorente, K. Bundy, C. Choi, H. Jeong,
H. Korhonen, P. Kubánek, J. Lim, A. Moskvitin,
T. Muñoz-Darias, S. Pak & I. Parrish

Nature **480**, 72–74 (2011).

The Acknowledgements of this Letter should have included the following line: “We thank A. J. Castro-Tirado for help in obtaining the data from the BTA 6-m telescope and comments on an early draft.” This has been added to the PDF and HTML versions online.

CAREERS

TURNING POINT Bioinformatician's career benefits from good timing **p.123**

NON-PROFIT SECTOR Despite job growth, research posts are under pressure **p.123**

NATUREJOBS For the latest career listings and advice www.naturejobs.com



COLUMN

Global collaboration

Mastering scientific protocols is only half the battle when it comes to conducting a clinical study in another country, says **Andrew Fung**.

In 2010, I began work on a project to create a new type of malaria test. I wanted to diagnose the disease using saliva, with an inexpensive test that looked, felt and tasted like chewing gum, rather than through the conventional blood sample. I had first proposed the idea in an effort to impress at an interview for a postdoctoral position. To my great delight, it gained funding from the Grand Challenges Explorations programme of the Bill & Melinda Gates Foundation in Seattle, Washington. I was aiming to land a job designing medical devices, but ended up with a grant for a blue-skies idea.

Had I bitten off more than I could chew? I was a postdoc at the University of California, Los Angeles (UCLA), at the time, and Los Angeles has few malaria patients. How would I confirm the prevalence of the malaria-causing parasite's proteins in saliva? I would have to go overseas and, crucially, I would have to build

a collaborative group of experts whose skills complemented my own.

Successfully managing this collaborative, international clinical study required not only the right team, but also careful planning, dealing with regulations and adapting to unexpected circumstances in a different culture. I learned that it takes more than a creative idea and bench experience to guide a project to its intended destination.

ASSEMBLE RESEARCH ALLIES

I had training in protein detection, an essential skill for analysing samples in the field, but needed help to run a clinical study in a malaria-endemic region. A friend and mentor — Theodore Moore, the clinical director of the Pediatric Blood and Marrow Transplant Program at UCLA — shared my interest in making new medical tools for developing communities.

In 1989, as a senior medical student, Moore had started to volunteer for relief efforts around the world. He ended up training medical professionals in countries from southern Africa to the Middle East. I wanted Moore on the team because of his professional expertise and his sympathy for the cause.

I also needed access to malaria patients, and a partner to handle field operations. During Moore's medical missions in the Philippines, he had befriended just the right person: Daniel Horton, then administrator of the Palawan Baptist Hospital (PBH), which he had helped to establish. During his 36 years in rural Palawan, the province with the highest incidence of malaria in the Philippines, Horton had seen more malaria patients than he could remember. He knew how to improvise under unexpected conditions, and could help us to navigate the local culture. Moore and Horton had ►

BELLE MELLOR

► strengths in areas that were my weaknesses.

At our kick-off meeting, we agreed on the objectives and timeline of our work, drawing up an overall strategy and specific contingency plans. This is the best time to flag up foreseeable disruptions, address potential ethical problems and agree on how to share data and intellectual property. In our case, setting a general order of authorship before pouring effort into the collaborative melting pot gave us a smooth working relationship, free from surprise competition.

My advice is to make sure that the group meets in person before you start, especially if lead members haven't worked together before. Moore was familiar with the PBH and UCLA and coordinated all our efforts, but I didn't meet Horton or his team until we arrived in Palawan for the fieldwork. A pre-emptive visit to the hospital would have helped us to spot some lurking problems with site logistics and the availability of lab ware. Teleconferences should be saved for short communications and status updates.

ADAPT YOUR MATERIALS AND METHODS

When researchers are working internationally, experimental methods need to be adapted to the infrastructure of the host facility. For example, our original protocols recommended storing saliva at -80°C . But ultra-low-temperature freezers weren't an option at the PBH, which has public electricity for just half the day, so we redesigned our approach to use a domestic freezer. Horton's researchers at the PBH also refitted their lab with a combination of electrical outlets at 120 volts and 220 volts so that we could use equipment from the United States in their hospital.

The climate affected our materials and processes. Reactions at room temperature have an entirely different behaviour in a tropical environment at 34°C than in one where the normal temperature is around 25°C . By using dry reagents stored in a desiccator, we kept protocols portable and stable. Because we would eventually transfer those protocols to Horton's team, we got other operators to reproduce our experiments in a few different labs, to anticipate possible hiccups.

Some basic materials may be difficult to find abroad. The hospital ended up stocking 40 litres of distilled water just for our chemicals, much more than they typically kept on hand. Will your reagents need a low-temperature supply chain? Team members who flew from Los Angeles carried perishable reagents in temperature-controlled boxes in checked baggage. At the PBH, we crushed ice to keep our working samples chilled using a domestic appliance for making halo-halo, a local fruit dessert.

Timing and logistics were crucial. Although

my grant lasted for one year, the seasonal nature of malaria gave us only three months in which to collect samples. Early on, we ruled out shipping specimens from the Philippines to the United States for analysis, because of the high cost and legal barriers. That meant we had to transport chemicals and instruments overseas for fieldwork. Whether you choose to travel with these materials as passenger baggage or to ship them ahead as cargo, arranging their passage can be less than straightforward.

Customs carnets can be very convenient documents for arranging temporary import of scientific equipment: they act as merchandise passports, allowing simplified, tax-free customs procedures. Unfortunately, many countries, including the Philippines, do not accept carnets.

To find the best alternative, we sought guidance from a logistics officer for the International Medical Corps, a global organization for humanitarian relief based in Santa Monica, California. He made sure that we prepared meticulous packing lists and an official letter of intent designating our materials solely for academic research. By being detailed about what we were shipping and clear about our purpose, we could apply for authorization to import before we arranged for transit.

Because the PBH was not a licensed importer, we had to work with a door-to-door international forwarder and its clearing agent in Manila. The agent caught us off guard with a very high and unjustified quote for customs handling and brokerage. But with help from a friend who was chief customs officer in Manila, we ended up with a proper quote at a fraction of the original fee, and a personal escort through airport customs. Whatever import method you choose, you should make sure that you know what to expect at customs clearance in every country you have to go through. To avoid falling victim to price gouging, get quotes from more than one vendor. If anything seems grossly out of place, someone is probably not complying with local laws.

REGULATORY CHALLENGES

Logistics will get your boxes onto site, but the work that the scientists do there will be governed by myriad regulations. International collaborations involving clinical medical research require oversight from multiple institutional review boards (IRBs) charged with the protection of human subjects. Each review board has specific requirements for the protocol and consent forms.

To learn how to get through this often long and iterative process of approval most efficiently, Moore met with representatives from the IRBs of both UCLA and the PBH, as well as other experienced investigators. They explained best practice for international collaborations — cooperating with local government, upholding the institutions' integrity, writing protocols to prioritize patient safety

— and aspects unique to UCLA. By working with the decision-makers from the start, we got protocols approved on the first pass.

Sometimes institutional approval isn't enough. Make sure to include local-government officials in your regulatory team, to head off authorities who may scrutinize your project for financial gain. As we prepared to enrol our first patients at the PBH, a member of a Philippines environmental and cultural heritage group told Horton that the study needed its permission to proceed, implying that a bribe

of some sort was necessary. We felt our collective arm being twisted.

Moore had already sent information about our study to Abraham Mitra, the governor of Palawan.

When he and Horton sat down in Mitra's office and explained our circumstances, the governor was taken aback. "This is about malaria," he said. "What does it have to do with the environment?" A phone call was made and the problem evaporated. With written endorsement from the Provincial Health Office, we were clear of further regulatory obstacles. It never hurts to have friends inside the system.

WATCH OUT FOR UNEXPECTED TWISTS

In research, as in health, remember that an ounce of prevention is worth a pound of cure. Nevertheless, no amount of risk management — whether pertaining to technical uncertainty, complex logistics or impenetrable regulations — guarantees a smooth operation. We experienced some extra bumps first hand.

The oil company Pilipinas Shell Petroleum has operations in Palawan, and its social-development arm had partnered with the provincial government from 1999 onwards to launch a major anti-malaria campaign. Just before our study, their programme increased insecticide spraying, distributed bed nets and opened a new central clinic not far from the PBH. The timing of these interventions markedly reduced the number of malaria cases seen at the PBH, which could have jeopardized the project. We had to extend the enrolment period by several months to get enough people for our pilot study, and still ended up with fewer than we had expected.

International partnerships are as much about people as they are about science. They open unique avenues for discovery by empowering scientists to explore beyond their accustomed borders. So above all, enjoy the adventure. Granted, it is a risky enterprise, but one full of professional reward and personal satisfaction. ■

Andrew Fung is a postdoc in the electrical and computer engineering department at the University of Waterloo in Ontario, Canada.

"International partnerships are as much about people as they are about science."

TURNING POINT

Ewan Birney

EMBL-EBI

Ewan Birney helped to assemble the genetic database for the international Human Genome Project. In April, he will become an associate director at the European Molecular Biology Laboratory's European Bioinformatics Institute (EBI) in Hinxton, UK, where he has worked since 2000.

How did you launch your scientific career?

I was always drawn to biology. I went to Eton College, UK, a secondary school that has links with Cold Spring Harbor Laboratory (CSHL) in New York. Between school and university, it is traditional in England to take a 'gap year'. So I did an internship at the CSHL, which had a big effect on me. Working at an institution headed by James Watson, one of the key people involved in deciphering the structure of DNA, was probably one of the biggest reasons I went into science.

Describe the most significant turning point in your career.

I was a graduate student at the Wellcome Trust Sanger Institute in Hinxton, and was involved in the Human Genome Project from 1999 to 2000, when everything went crazy because of the competition between public and private efforts to sequence the genome. There was a constant push to release results. The Sanger was a major sequencing centre, and I had to grow up scientifically very quickly — learning to solve problems and get results — on account of the pace of discovery and the constant changes in technology. I effectively skipped the postdoc phase. I am sure I will never experience an atmosphere like that again. It was intensely demanding: every two weeks the team needed to create a new tool, or provide key information to show that we were on top of things.

What was your biggest accomplishment there?

In the final push of the Human Genome Project in 2000, we created Ensembl, a web browser for genetic information. We worked like idiots for two years to develop it. A huge amount of what we did fell on a few people who could understand databases and bioinformatics. Ensembl now holds genetic information from more than 60 species and a lot of tools for analysis — all freely available.

What will be your managerial focus at the EBI?

In 2006 I handed off Ensembl, because at some point, you have to trust people to improve what you've begun. It is wrenching to let go of something you've helped to create.



But you have to if you want it to live. That requires a lot of coaching and mentoring of those taking over the project. Hiring excellent people, setting up situations for them to succeed in and coaching and mentoring them will be key parts of my new position.

How did you come to bioinformatics?

Bioinformatics is just a word to describe part of modern data-intensive molecular biology. Ever since a senior scientist advised me to learn the C programming language when I was at the CSHL, I've used whatever methods or tools I can find. Very early on, I adapted a programme to scan databases of expressed sequence tags, short bits of DNA that help to identify genes. I continue to teach myself programming and statistics. In general, I think about what problems I would like to address, what theories are out there and what computational or mathematical approaches might work. For instance, some colleagues mentioned that combinatorial mathematics might be a good way to understand genome assembly. At first I didn't understand how, but I finally got it, then went back to the maths textbooks and refined my approach.

What's the secret of your success?

I just really enjoy doing science. And that has helped me to get through some difficult times, when I've pushed myself and others perhaps too hard. The other thing that leads to success is trusting collaborators and the people you hire; I think that, often, we don't put enough trust in the scientists around us and it hinders progress. I am very lucky to be a part of the EBI and to be moving into such a central role. I enjoy nearly everything about it — although there are one or two meetings I could do without. ■

INTERVIEW BY PAUL SMAGLIK

CANADA

Pension values shrink

Some Canadian academics will have to increase contributions to their pension plans as a consequence of the recession, says the Canadian Association of University Teachers (CAUT) in Ottawa, which represents faculty members at about 70 public institutions. Neil Tudiver, CAUT's assistant executive director of collective bargaining, says that the economic downturn has decreased the value of faculty pension accounts. Provincial regulations call for reductions to public universities' costs, forcing institutions to consider plan adjustments to make up the shortfall. Tudiver says that pension contributions for affected plans will rise by 2% on average.

NON-PROFIT SECTOR

Uncertain prospects

Tightening of federal support for US science is creating a mixed employment outlook for academic researchers, says the author of *Holding the Fort: Nonprofit Employment during a Decade of Turmoil*. The report, released on 18 January, found that employment in the US non-profit sector grew by an average of 2.1% each year from 2000 to 2010, despite the recession; growth in education alone was 2.6%. Although universities and health care saw some of the biggest job increases, a slowdown in federal funding is placing a strain on academic research, suggests Lester Salamon, a policy expert at Johns Hopkins University in Baltimore, Maryland, and lead author of the report. Salamon predicts a shift to private sources of research funding, including industry.

UNITED STATES

Postdoc stipend grows

For the second year running, postdocs funded through US National Institutes of Health (NIH) traineeships will receive a 2% stipend increase. In 2011, 6,686 postdocs received the stipend. The NIH announced the increase on 20 January for trainees — including undergraduates, PhD students and postdocs — who receive the Ruth L. Kirschstein National Research Service Award (NRSA). Entry-level NRSA stipends still fall short of the US\$45,000 advocated by the US National Postdoc Association (NPA) in Washington DC; last year, an informal NPA survey of 74 US institutions found that 35 of them base their postdoc salaries on the NRSA, which increased by 1% in 2010.

WORDS AND MUSIC

A performance of note.

BY RONALD D. FERGUSON

Yeah, you got me, fair and square. I didn't think you'd recognize me. I am the government translator, the guy who lurks off the shoulder of Space Systems' chief negotiator and whispers connotations, corrections and culpability in case the computer renders too literal a verdict on Utmano phrasing. I've got nothing to do with speakos. Talk with the programmers for input-output problems.

I hope you're recording this, because it will be my only interview. Please, just a few questions, so quit yelling. I'll tell the story, and then I go home. Understood?

Here's what happened. The meeting began with the formal greeting...

Example? You want to know what the Utmano said? Do you speak Utmano? Oh, the literal translation into English? Right. Let me see...

It is beautiful weather we shall be having tomorrow wasn't it?

That's not the literal translation of the Utmano greeting, but that's as close as I'm likely to get on a Tuesday. Today is Wednesday? No, I can't do better today. Of course, the literal translation isn't what the Utmano meant. From the variations of pitch within the context of the meeting, my best interpretation is:

We should complete negotiations successfully tomorrow based on the current tolerant atmosphere and previously adopted ground rules.

The computer's translation was close, possibly better based on pitch and less accurate where context was important. Diphthong? No pitch, like I said. Intonation would be a reasonable description. Thai is the human intonation language that comes to mind, but Utmano is like intonation on steroids. Spoken Utmano must be sung.

No, I don't think an Utmano looks like a whale. That's uncalled for.

Speak Utmano? Unaided, I cannot speak Utmano. My base vocabulary is okay, and I might squeeze out a few words, but the Utmano tonal range is 14 octaves. They have some vocabulary modifiers well above 20,000 hertz, which is likely to make your neighbour's dog bark. Humans have a more limited range. I use a synthesizer to...

Human range? Oh, I don't know. If you watch old videos, Julie Andrews had a four-octave range, Mariah Carey perhaps five, but most people can't do that. I certainly can't.

I'm good at listening to the symphony that Utmano call speech. I'm not Mozart, but I have a good ear, and training as a professional musician.

No, Julie Andrews. You know, *The Sound of Music*. Okay?

Symphony? Each Utmano has a double set of vocal chords and can sound two notes at once. An Utmano can sing harmony with itself. The Utmano divide an octave into 32 parts. That's why we rely on computers and Fourier analysis...



No. I understand your question, but Utmano translation doesn't compare well to translations among human languages. The Utmano language has a peculiar view of tense, you know, past, present, future, in its sentences — well maybe not sentences, but the complete-thought communication structure. Psychologists claim that the Utmano have a lingering, vivid, recent memory combined with a mild prescience that blends with their perception of the 'now'. It sounds like gobbledygook, but they claim that the Utmano idea of the present spans from the middle of last week to a couple of hours from now.

In translating from English to Greek, or French, or whatever Earth language you

ON NATURE.COM
Follow Futures on
Facebook at:
go.nature.com/mtoodm

pick, we have the advantage of common human experiences. Humans share uncommonly little with the Utmano.

No. I'm not making excuses. I accept the blame for missing any subtle nuance, but the Utmano response was undocumented. Google the literature, and the only violent reactions are for personal insults...

I'll stick to the topic. I repeat for the record. Intonation is not a critical characteristic of most human languages, but to the Utmano, a change in pitch is an essential communication detail.

Look, I've got to go. I need to finish this interview. I can't give any details of the ongoing negotiations, because those are secret. I can confirm the official report. Yes, I know the report is short. Do you have a copy? Okay, I'll read it for those in the back.

"Chief Negotiator Simon Mann died today when he bade farewell to the Utmano Ambassador. A focused eruption of sonic energy homogenized Dr Mann's brain. With no hope for recovery, his family directed he be removed from life support."

You heard that the sonic energy was a blast of rage from the Utmano Ambassador? No. I cannot confirm that. Yes, I was there, but appearances aren't always reality.

Yes, I was told that Dr Mann practised a few phrases in Utmano. No, I didn't coach him. I was never asked to coach him. Besides, I use the very tedious process of keyboarded voice modulation when I speak Utmano. Dr Mann wasn't a musician, and I wouldn't know how to coach him. I certainly didn't know that Dr Mann planned to speak in Utmano or I would have advised against it.

What? John Kennedy at the Berlin Wall, the '*Ich bin ein Berliner*' speech? Yeah, I've seen the newsreel. No, I don't think Kennedy said 'I'm a jelly doughnut.' No, I don't know whether Dr Mann planned to run for elected office.

Negotiations? Negotiations will resume in two weeks, once we are well past the Utmano concept of the present. No, I don't believe that is 'pout time' or 'holding a grudge'.

One last question? Okay, you, in the back. Did Dr Mann make a mistake? All I've got is my humble opinion, not as a translator, but as a musician.

Like many tone-deaf amateurs, he thought he could sing. ■

Ronald D. Ferguson has decided that writing fiction is more fun than writing college textbooks. He lives with his wife and five feral cats on two acres of the Texas Hill Country.

Enhancer decommissioning by LSD1 during embryonic stem cell differentiation

Warren A. Whyte^{1,2*}, Steve Bilodeau^{1*}, David A. Orlando¹, Heather A. Hoke^{1,2}, Garrett M. Frampton^{1,2}, Charles T. Foster^{3,4}, Shaun M. Cowley⁴ & Richard A. Young^{1,2}

Transcription factors and chromatin modifiers are important in the programming and reprogramming of cellular states during development^{1,2}. Transcription factors bind to enhancer elements and recruit coactivators and chromatin-modifying enzymes to facilitate transcription initiation^{3,4}. During differentiation a subset of these enhancers must be silenced, but the mechanisms underlying enhancer silencing are poorly understood. Here we show that the histone demethylase lysine-specific demethylase 1 (LSD1; ref. 5), which demethylates histone H3 on Lys 4 or Lys 9 (H3K4/K9), is essential in decommissioning enhancers during the differentiation of mouse embryonic stem cells (ESCs). LSD1 occupies enhancers of active genes that are critical for control of the state of ESCs. However, LSD1 is not essential for the maintenance of ESC identity. Instead, ESCs lacking LSD1 activity fail to differentiate fully, and ESC-specific enhancers fail to undergo the histone demethylation events associated with differentiation. At active enhancers, LSD1 is a component of the NuRD (nucleosome remodelling and histone deacetylase) complex, which contains additional subunits that are necessary for ESC differentiation. We propose that the LSD1–NuRD complex decommissions enhancers of the pluripotency program during differentiation, which is essential for the complete shutdown of the ESC gene expression program and the transition to new cell states.

The histone H3K4/K9 demethylase LSD1 (also known as KDM1A) is one of the chromatin regulators that have been implicated in the control of early embryogenesis^{6–8}. Loss of LSD1 leads to embryonic lethality, and ESCs lacking LSD1 function fail to differentiate into embryoid bodies^{6–8}. These results suggest that LSD1 contributes to changes in chromatin that are critical to the differentiation of ESCs, but the role of LSD1 in this process is not yet understood. To investigate the function of LSD1 in ESCs, we first identified the sites it occupies in the genome by using chromatin immunoprecipitation coupled with massively parallel DNA sequencing (ChIP-Seq; Fig. 1 and Supplementary Fig. 1). The results revealed that LSD1 occupies the enhancers and core promoters of a substantial population of actively transcribed and bivalent genes (Fig. 1a, b and Supplementary Table 1). Inspection of individual gene tracks showed that LSD1 occupies well-characterized enhancer regions together with the ESC master transcription factors Oct4, Sox2 and Nanog and the Mediator coactivator (Fig. 1b and Supplementary Fig. 1). Loci bound by Oct4, Sox2 and Nanog are generally associated with Mediator and p300 coactivators and have enhancer activity^{9,10}. A global view of enhancer regions occupied by Oct4, Sox2, Nanog and Mediator confirmed that 97% of the 3,838 high-confidence enhancers were also occupied by LSD1 ($P < 10^{-9}$) (Fig. 1c and Supplementary Table 2). This is consistent with evidence that LSD1 can interact with Oct4 (refs 11, 12). LSD1 signals were also observed at core promoter regions with RNA polymerase II (Pol II) and TATA-binding protein (TBP; Fig. 1d). The

density of LSD1 signals at enhancers was higher than at core promoters ($P < 10^{-16}$; Supplementary Fig. 1), indicating that LSD1 is associated predominantly with the enhancers of actively transcribed genes in ESCs.

It was striking to find that LSD1 is associated with active genes in ESCs because previous studies have shown that LSD1 is not essential for the maintenance of ESC state but is required for normal differentiation^{6–8}. We used an ESC differentiation assay to further investigate the involvement of LSD1 in cell state transitions (Fig. 2a, b). Prolonged depletion of Oct4 in ZHBTc4 ESCs with doxycycline causes loss of pluripotency and differentiation into trophectoderm¹³. As expected, loss of Oct4 expression led to a rapid loss of ESC morphology and a marked decrease in the levels of SSEA-1 and alkaline phosphatase, two markers of ESCs (Fig. 2c and Supplementary Fig. 2). When these ESCs were treated with the LSD1 inhibitor tranylcypromine (TCP) during Oct4 depletion, they failed to undergo the morphological changes associated with differentiation of ESCs (Fig. 2c). Instead, the TCP-treated cells formed small colonies resembling those of untreated ESCs and maintained expression of SSEA-1 and alkaline phosphatase (Fig. 2c and Supplementary Fig. 3). Very similar results were partly obtained in LSD1 knockout ESCs (Supplementary Figs 4 and 5) and in cells treated with another LSD1 inhibitor, pargyline, or a short hairpin RNA against LSD1 (Supplementary Figs 2 and 3). LSD1 inhibition also caused an increase in cell death during differentiation, as has been observed with cells lacking LSD1 in other assays^{7,8}. These results suggest that LSD1 may be required for ESCs to silence the ESC gene expression program completely.

Further analysis of ESCs that were forced to differentiate in the absence of LSD1 activity confirmed that these cells failed to make a complete transition from the ESC gene expression program; although key genes of the trophectoderm gene expression program were activated, including *Cdx2* and *Esx1* (ref. 14), there was incomplete repression of many ESC genes, including *Sox2* and *Fbx15* (Fig. 2d). A global analysis confirmed that a set of genes neighbouring LSD1-occupied enhancers in ESCs are repressed during differentiation and that the repression of this set of genes is partly relieved in the presence of TCP (Fig. 2e and Supplementary Table 3). Similar results were obtained with LSD1 knockout cells (Supplementary Figs 4 and 5) and with cells treated with either pargyline or a short hairpin RNA against LSD1 (Supplementary Fig. 3). These results indicate that the trophectoderm differentiation program can be induced in cells lacking LSD1 function, but the ESC program is not fully silenced in these cells.

To gain further insight into the role of LSD1 in ESC differentiation, we investigated whether LSD1 is associated with previously described complexes, including NuRD, cofactor of REST (CoREST), and the androgen receptor/oestrogen receptor complexes^{8,15–17}. We first studied whether the LSD1 found at Oct4-occupied genes is a component of NuRD, because Oct4 and Nanog have been reported to interact with several components of NuRD^{11,12,18}. ChIP-Seq experiments

¹Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Department of Molecular Biology, Adolf-Butenandt Institut, Ludwig-Maximilians-Universität München, 80336 Munich, Germany. ⁴Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK.

*These authors contributed equally to this work.

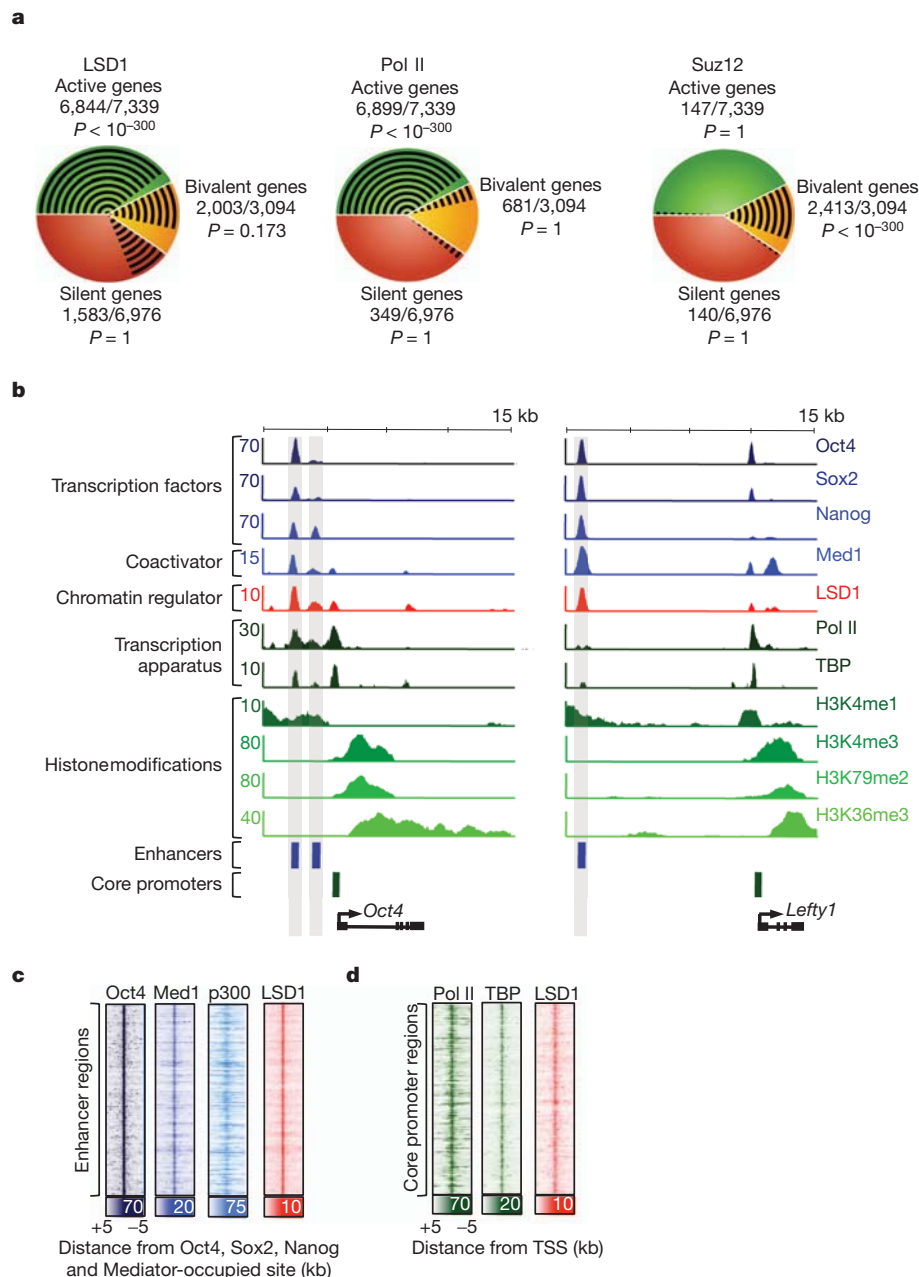


Figure 1 | LSD1 is associated with enhancer and core promoter regions of active genes in ESCs. **a**, LSD1 occupies a substantial population of actively transcribed genes in murine ESCs. The pie charts depict active (green), bivalent (yellow) and silent (red) genes, and the proportion (black lines) occupied by LSD1, Pol II or the Polycomb protein Suz12 (Supplementary Table 1 and Supplementary Information). The numbers of genes bound and the total number of genes in each of the active, bivalent and silent classes are shown. LSD1 ChIP-Seq data are from combined biological replicates using an antibody specific for LSD1 as determined by knockdown experiments (Supplementary Fig. 1). The P value for each category was determined by a hypergeometric test. **b**, LSD1 occupies enhancers and core promoter regions of actively transcribed genes. Shown are ChIP-Seq binding profiles (reads per million) for ESC transcription factors (Oct4, Sox2, Nanog), coactivator (Med1), chromatin regulator (LSD1), the transcriptional apparatus (Pol II, TBP) and histone modifications (H3K4me1, H3K4me3, H3K79me2, H3K36me3) at the *Oct4*

(*Pou5f1*) and *Lefty1* loci in ESCs, with the y-axis floor set to 1. Gene models and previously described enhancer regions^{27–29} are shown below the binding profiles. **c**, LSD1 occupies enhancer sites. A density map is shown of ChIP-Seq data at Oct4, Sox2, Nanog and Med1 co-occupied enhancer regions. Data are shown for an ESC transcription factor (Oct4), coactivators (Med1 and p300) and a chromatin regulator (LSD1) in ESCs. Enhancers were defined as Oct4, Sox2, Nanog and Mediator co-occupied regions. More than 96% of the 3,838 high-confidence enhancers were co-occupied by LSD1 ($P < 10^{-9}$). Colour scale indicates ChIP-seq signal in reads per million. **d**, LSD1 occupies core promoter sites. Shown is a density map of ChIP-Seq data at transcriptional start sites (TSSs) of genes neighbouring the 3,838 previously defined enhancers (c). Data are shown for components of the transcription apparatus (Pol II and TBP) and the chromatin regulator LSD1 in ESCs. Core promoters were defined as the closest TSS from each enhancer. Colour scale indicates ChIP-Seq signal in reads per million.

confirmed that NuRD subunits Mi-2 β , HDAC1 and HDAC2 together occupy sites with LSD1 at enhancers ($P < 10^{-9}$; Fig. 3 and Supplementary Table 1). Immunoprecipitation of LSD1 confirmed its association with Mi-2 β , HDAC1 and HDAC2 (Fig. 3b, c). We then investigated whether LSD1 is associated with CoREST; ChIP-Seq data revealed that

a minor fraction of LSD1 occupies sites together with CoREST and REST (2% and 6%, respectively) (Supplementary Fig. 6 and Supplementary Table 1). As expected, LSD1–REST sites were frequently found associated with neuronal genes (Supplementary Fig. 7 and Supplementary Table 4). Immunoprecipitation experiments confirmed that LSD1 is

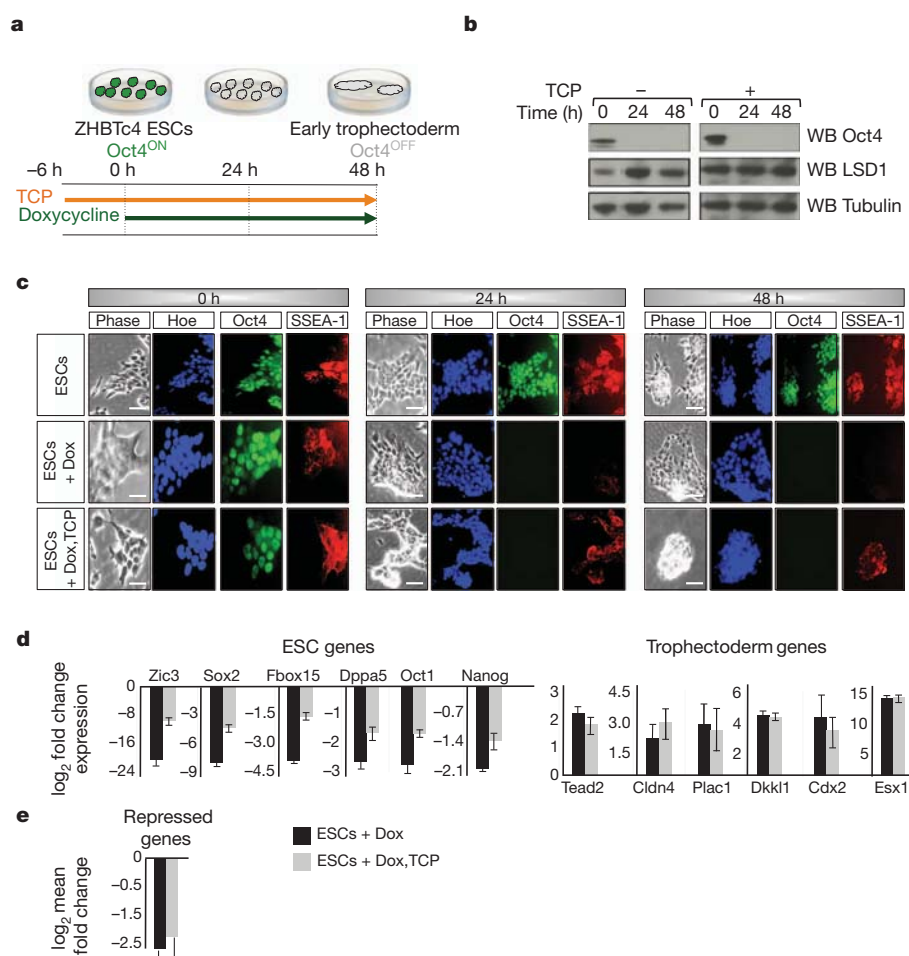


Figure 2 | LSD1 inhibition results in incomplete silencing of ESC genes during differentiation. **a**, Schematic representation of trophoblast differentiation assay using the doxycycline-inducible Oct4 shutdown murine ESC line ZHBTc4. Treatment with doxycycline for 48 h leads to depletion of Oct4 and early trophoblast specification. Cells were treated with dimethylsulphoxide (DMSO; control) or the LSD1 inhibitor TCP for 6 h before 2 $\mu\text{g ml}^{-1}$ doxycycline was added for a further 24 or 48 h. **b**, Treatment of ZHBTc4 ESCs with doxycycline leads to loss of Oct4 proteins. Oct4 and LSD1 protein levels in nuclear extracts determined by western blotting (WB) before and after treatment of ZHBTc4 ESCs with 2 $\mu\text{g ml}^{-1}$ doxycycline. Tubulin served as loading control. **c**, Doxycycline (Dox)-treated cells treated with TCP maintained SSEA-1 cell surface marker expression. Cells were stained for DNA

(Hoechst; Hoe), Oct4 and SSEA-1. Scale bar, 100 μm . **d**, Expression of selected ESC and trophoblast genes 48 h after Oct4 depletion in dimethylsulphoxide-treated and TCP-treated cells (black and grey bars, respectively). Treatment with TCP partly relieved repression of ESC genes but did not affect upregulation of trophoblast genes. Error bars show s.d. from biological replicates. **e**, Genes neighbouring LSD1-occupied enhancers are less downregulated during ESC differentiation after TCP treatment. Shown is the mean fold change in expression of the 630 downregulated (at least 1.25-fold; $P < 0.01$) genes nearest LSD1-occupied enhancers (Fig. 1c) during differentiation of TCP-treated and untreated control cells. Alleviation of repression is significantly higher (asterisk, $P < 0.005$) for LSD1 enhancer-bound repressed genes than for all repressed genes.

associated with CoREST (Fig. 3b, c). Androgen receptor and oestrogen receptor are not expressed in ESCs, as indicated by the lack of histone H3K79me2 and H3K36me3 (modifications associated with transcriptional elongation) at the genes encoding these proteins (Supplementary Table 1). Further examination of the ChIP-Seq data revealed that enhancers were significantly more likely to be occupied by the LSD1 and NuRD proteins than by REST and CoREST ($P < 10^{-9}$) (Fig. 3d and Supplementary Fig. 8). Multiple components of NuRD are dispensable for ESC state but are required for normal differentiation^{6,19–21}. ESCs with decreased levels of the core NuRD ATPase Mi-2 β failed to differentiate properly and partly maintained expression of SSEA-1, alkaline phosphatase and ESC genes (Supplementary Fig. 9), which are the same phenotypes as those we observed with decreased levels of LSD1. These results indicate that LSD1 at enhancers is associated with a NuRD complex that is essential for normal cell state transitions.

Nucleosomes with histone H3K4me1 are commonly found at enhancers of active genes and are a substrate for LSD1 (refs 5, 22). If LSD1-dependent H3K4me1 demethylase activity is involved in enhancer silencing during ESC differentiation, LSD1 inhibition should cause the

retention of H3K4me1 levels at active ESC enhancers when differentiation is induced. During trophoblast differentiation with control ESCs, we found decreased levels of p300 and H3K27ac at a set of active ESC enhancers, suggesting that these enhancers were being silenced (Supplementary Fig. 10). The levels of H3K4me1 at enhancers were also decreased, as seen for example at *Lefty1* (Fig. 4a and Supplementary Table 5), whereas the levels of H3K4me1 increased at newly active trophoblast genes such as *Gata2* (Fig. 4b). In contrast, H3K4me1 signals were higher at LSD1-occupied enhancers in differentiating ESCs treated with TCP than in control cells, including *Lefty1* and *Sox2* (Fig. 4a, c). Most enhancers (1,722 of 2,755) that were occupied by LSD1 and that experienced decreased levels of H3K4me1 during differentiation retained H3K4me1 in TCP-treated ESCs, in contrast to untreated control differentiating ESCs (Fig. 4d, e). These results are consistent with the model that LSD1 demethylates H3K4me1 at the enhancers of ESC-specific genes during differentiation and that this activity is essential to fully repress the genes associated with these enhancers.

Our results indicate that an LSD1–NuRD complex is required for silencing of ESC enhancers during differentiation, which is essential

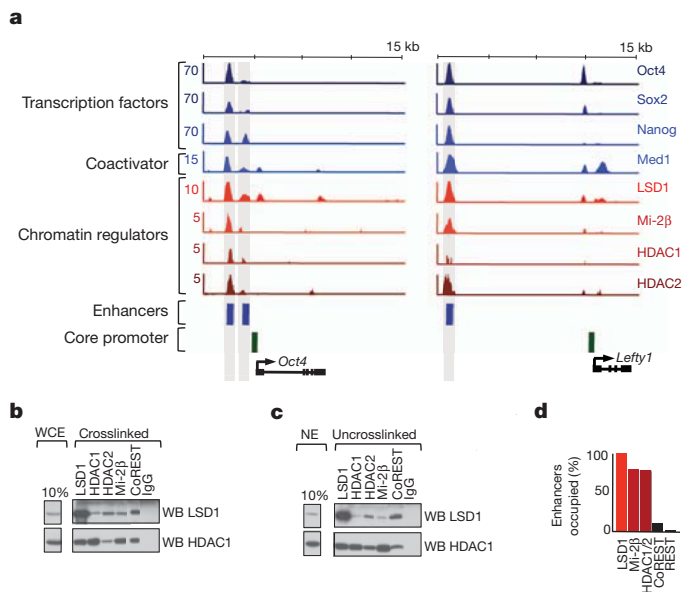


Figure 3 | LSD1 is associated with a NuRD complex at active enhancers in ESCs. **a**, NuRD components occupy enhancers and core promoter regions of actively transcribed genes. Shown are ChIP-Seq binding profiles (reads per million) for transcription factors (Oct4, Sox2, Nanog), coactivator (Med1) and chromatin regulators (LSD1, Mi-2 β , HDAC1, HDAC2), at the *Oct4* (*Pou5f1*) and *Lefty1* loci in ESCs, with the y-axis floor set to 1. Gene models and previously described enhancer regions^{27–29} are depicted below the binding profiles. **b**, LSD1 is associated with NuRD components Mi-2 β , HDAC1 and HDAC2, as well as with CoREST. LSD1 and HDAC1 are detected by western blotting (WB) after immunoprecipitation of crosslinked whole cell extract (WCE) with anti-LSD1, anti-HDAC1, anti-HDAC2, anti-Mi-2 β or anti-CoREST antibodies. IgG is shown as a control. **c**, LSD1 and HDAC1 are detected by western blotting after immunoprecipitation of uncrosslinked nuclear extracts (NE) using anti-LSD1, anti-HDAC1, anti-HDAC2, anti-Mi-2 β or anti-CoREST antibodies. IgG is shown as a control. **d**, The occupancy of enhancers by NuRD proteins (Mi-2 β , HDAC1 and HDAC2) is significantly greater than the occupancy by CoREST or REST ($P < 10^{-9}$). The height of the bars represents the percentage of the 3,838 enhancers co-occupied by LSD1, NuRD proteins (Mi-2 β , and either HDAC1 or HDAC2), CoREST and REST.

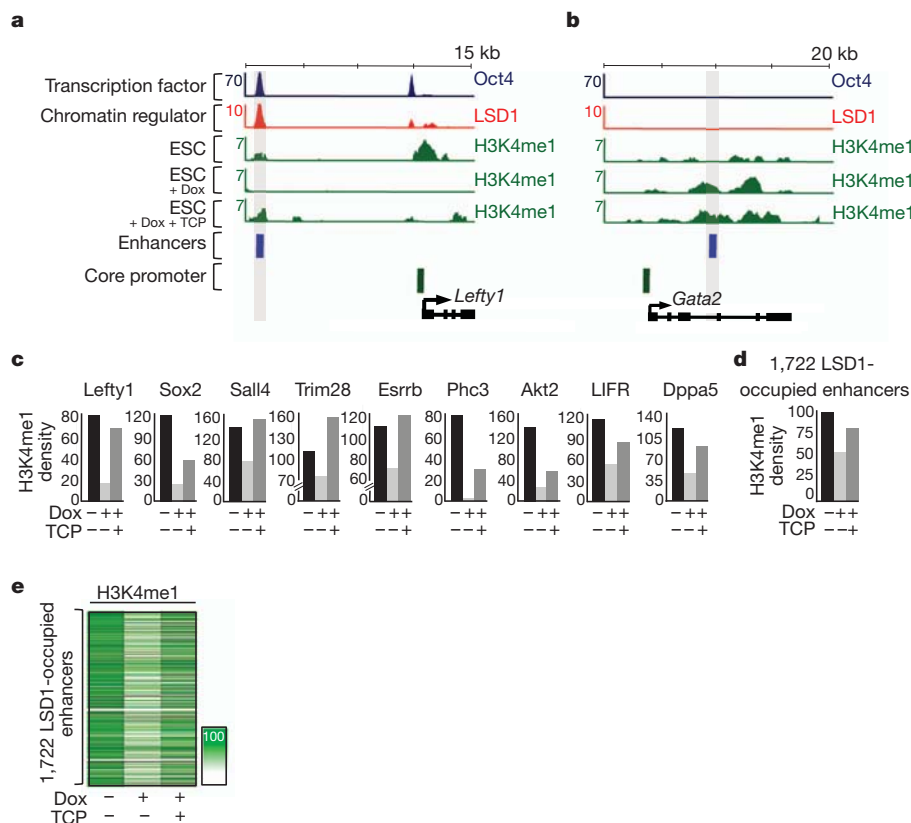


Figure 4 | LSD1 is required for H3K4me1 removal at ESC enhancers.

a, H3K4me1 levels are decreased at LSD1-occupied enhancers during ESC differentiation, and this effect is partly blocked on treatment with TCP. Dox, doxycycline. **b**, Treatment with TCP does not affect the increase in H3K4me1 levels at trophoctodermal genes during differentiation. Shown are ChIP-Seq binding profiles (reads per million) for Oct4 and LSD1 at the *Lefty1* and *Gata2* loci in ESCs. Below these profiles, histone H3K4me1 levels are shown for ZHBTc4 control ESCs, cells treated with doxycycline for 48 h to repress Oct4 and induce differentiation (ESCs + Dox), and ESCs treated with doxycycline and TCP (ESCs + Dox + TCP). For appropriate normalization, ChIP-Seq data for histone H3K4me1 is shown as rank normalized reads per million with the y-axis floor set to 1 (Supplementary Information). Gene models and previously described enhancer regions^{29,30} are depicted below the binding profiles. **c**, Sum

of the normalized H3K4me1 density ± 250 nucleotides surrounding LSD1-occupied enhancer regions before and during trophoctoderm differentiation in the presence or absence of TCP. The associated genes were identified on the basis of their proximity to the LSD1-occupied enhancers. **d**, Sum of the normalized H3K4me1 density ± 250 nucleotides surrounding 1,722 LSD1-occupied enhancers before and during differentiation in the presence or absence of TCP. Of the 2,755 LSD1-occupied enhancers with decreased levels of H3K4me1 on differentiation, 63% (1,722) had higher H3K4me1 levels after TCP treatment ($P < 10^{-16}$). **e**, Heat map displaying the sum of the normalized H3K4me1 density ± 250 nucleotides surrounding the 1,722 LSD1-occupied enhancers that retained H3K4me1 in TCP-treated ESCs compared with untreated control differentiating ESCs. Colour scale indicates ChIP-Seq signal in normalized reads per million.

for complete shutdown of the ESC gene expression program and the transition to new cell states. These results, together with those of previous studies on NuRD function^{18,21,23,24}, suggest the following model for LSD1–NuRD in enhancer decommissioning. LSD1–NuRD complexes occupy Oct4-regulated active enhancers in ESCs but do not substantially demethylate histone H3K4 because the H3K4 demethylase activity of LSD1 is inhibited in the presence of acetylated histones^{23,24}. Enhancers occupied by Oct4, Sox2 and Nanog are also occupied by the HAT p300 and nucleosomes with acetylated histones (Supplementary Fig. 10)¹⁰. Thus, as long as the enhancer-bound transcription factors recruit HATs to enhancers, the net effect of having both HATs and NuRD-associated HDACs present is to have sufficient levels of acetylated histones to suppress LSD1 demethylase activity. During ESC differentiation, the levels of Oct4 and p300 are decreased, thus decreasing the level of acetylated histones, which in turn permits the demethylation of H3K4 by LSD1. Consistent with this model, we find that the shutdown of Oct4 leads to decreased levels of p300 and histone H3K27ac at enhancers that are occupied by Oct4 and LSD1 (Supplementary Figs 10 and 11), and this is coincident with decreased levels of methylated H3K4 (Fig. 4 and Supplementary Figs 12 and 13). This model would explain why key components of LSD1–NuRD complexes are not essential for the maintenance of ESC state but are essential for normal differentiation, when the active enhancers must be silenced. Additional HATs expressed in ESCs may also contribute to the dynamic balance of nucleosome acetylation. Future biochemical analysis of HAT, HDAC and demethylase complexes at enhancers will be valuable for testing this model and for further understanding how enhancers are regulated during differentiation.

We conclude that LSD1–NuRD complexes present at active promoters in ESCs are essential for normal differentiation, when the active enhancers must be silenced. Given that there is evidence that LSD1 is required for differentiation of multiple cell types^{6,25,26}, LSD1 is likely to be generally involved in enhancer silencing during differentiation. The ESC gene expression program can be maintained in the absence of many other chromatin regulators², and it is possible that some of these also have key functions in the transition from one transcriptional program to another during differentiation.

METHODS SUMMARY

ESC culture conditions. ESCs were grown on irradiated murine embryonic fibroblasts (MEFs) and passaged as described previously⁹. In drug treatment experiments, ESCs were split off MEFs and treated with 1 mM TCP or 3 mM pargyline to inhibit LSD1 activity. Lentiviral constructs were purchased from Open Biosystems and produced according to the *Trans*-lentiviral shRNA Packaging System (catalogue no. TLP4614).

Differentiation assay, immunofluorescence, and alkaline phosphatase staining. ZHBTc4 ESCs were split off MEFs in ESC medium containing 2 µg ml^{−1} doxycycline to decrease Oct4 expression levels. For immunofluorescence, ESCs were crosslinked, blocked and permeabilized before incubation with anti-Oct4 (Santa Cruz, sc-9081x; 1:200 dilution) or anti-SSEA1 (mc-480, Developmental Studies Hybridoma Bank; 1:20 dilution) antibodies. Alexa-conjugated secondary antibodies were used for detection. Staining of ESCs for alkaline phosphatase was achieved with the Alkaline Phosphatase Detection Kit (Millipore, SCR004). Cells were harvested at indicated time points for ChIP-Seq, quantitative polymerase chain reaction or expression array analyses.

ChIP-Seq. Chromatin immunoprecipitations (ChIPs) were performed and analysed as described previously⁹. The following antibodies were used: anti-LSD1 (Abcam, ab17721), anti-Mi-2b (Abcam, ab72418), anti-HDAC1 (Abcam, ab7028), anti-HDAC2 (Abcam, ab7029), anti-REST (Millipore, 07-579), anti-CoREST (Abcam, ab32631), anti-H3K4me1 (Abcam, ab8895), anti-p300 (Santa-Cruz, sc-584) and anti-H3K27Ac (Abcam, ab4729).

For ChIP-Seq analyses, reads were aligned with Bowtie and analysed as described in Supplementary Information.

Received 14 October 2010; accepted 16 December 2011.

Published online 1 February 2012.

1. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
2. Young, R. A. Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011).

3. Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192 (2009).
4. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
5. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
6. Wang, J. *et al.* Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature* **446**, 882–887 (2007).
7. Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature Genet.* **41**, 125–129 (2009).
8. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol. Cell. Biol.* **30**, 4851–4863 (2010).
9. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
10. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
11. Pardo, M. *et al.* An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**, 382–395 (2010).
12. van den Berg, D. L. *et al.* An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369–381 (2010).
13. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet.* **24**, 372–376 (2000).
14. Rossant, J. & Cross, J. C. Placental development: lessons from mouse mutants. *Nature Rev. Genet.* **2**, 538–548 (2001).
15. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).
16. Shi, Y. J. *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell* **19**, 857–864 (2005).
17. Wang, Y. *et al.* LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell* **138**, 660–672 (2009).
18. Liang, J. *et al.* Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature Cell Biol.* **10**, 731–739 (2008).
19. Dovey, O. M., Foster, C. T. & Cowley, S. M. Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proc. Natl Acad. Sci. USA* **107**, 8242–8247 (2010).
20. Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nature Cell Biol.* **8**, 285–292 (2006).
21. Scimone, M. L., Meisel, J. & Reddien, P. W. The Mi-2-like *Smed-Chd4* gene is required for stem cell differentiation in the planarian *Schmidtea mediterranea*. *Development* **137**, 1231–1241 (2010).
22. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39**, 311–318 (2007).
23. Forneris, F., Binda, C., Vanoni, M. A., Battaglioli, E. & Mattevi, A. Human histone demethylase LSD1 reads the histone code. *J. Biol. Chem.* **280**, 41360–41365 (2005).
24. Lee, M. G. *et al.* Functional interplay between histone demethylase and deacetylase enzymes. *Mol. Cell. Biol.* **26**, 6395–6402 (2006).
25. Musri, M. M. *et al.* Histone demethylase LSD1 regulates adipogenesis. *J. Biol. Chem.* **285**, 30034–30041 (2010).
26. Choi, J. *et al.* Histone demethylase LSD1 is required to induce skeletal muscle differentiation by regulating myogenic factors. *Biochem. Biophys. Res. Commun.* **401**, 327–332 (2010).
27. Okumura-Nakanishi, S., Saito, M., Niwa, H. & Ishikawa, F. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J. Biol. Chem.* **280**, 5307–5317 (2005).
28. Yeom, Y. I. *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894 (1996).
29. Nakatake, Y. *et al.* Klf4 cooperates with Oct3/4 and Sox2 to activate the *Lefty1* core promoter in embryonic stem cells. *Mol. Cell. Biol.* **26**, 7772–7782 (2006).
30. Ray, S. *et al.* Context-dependent function of regulatory elements and a switch in chromatin occupancy between GATA3 and GATA2 regulate Gata2 transcription during trophoblast differentiation. *J. Biol. Chem.* **284**, 4978–4988 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Lovén, M. H. Kagey, J. Downen, A. C. Mullen, A. Sigova, P. B. Rahl, T. Lee and members of Y. Shi's laboratory for experimental assistance, reagents and helpful discussions; and J.-A. Kwon, V. Dhanapal, J. Love, S. Gupta, T. Volkert, W. Salmon and N. Watson for assistance with ChIP-Seq, expression arrays and immunofluorescence imaging acquisition. This work was supported by a Canadian Institutes of Health Research Fellowship (S.B.), a Career Development Award from the Medical Research Council (S.M.C.), and by National Institutes of Health grants HG002668 and NS055923 (R.Y.).

Author Contributions W.A.W., S.B., H.A.H., C.T.F., S.M.C. and R.A.Y. designed, conducted and interpreted the ChIP-Seq, immunofluorescence and expression experiments. W.A.W., D.A.O. and G.M.F. performed data analysis. The manuscript was written by S.B., W.A.W., D.A.O., H.A.H., G.M.F. and R.A.Y.

Author Information ChIP-Seq and GeneChip expression data are deposited in the Gene Expression Omnibus under accession number GSE27844. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.A.Y. (young@wi.mit.edu).

Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity

Jorge Henao-Mejia^{1*}, Eran Elinav^{1*}, Chengcheng Jin^{1,2*}, Liming Hao³, Wajahat Z. Mehal⁴, Till Strowig¹, Christoph A. Thaiss¹, Andrew L. Kau^{5,6}, Stephanie C. Eisenbarth⁷, Michael J. Jurczak⁴, Joao-Paulo Camporez⁴, Gerald I. Shulman^{4,8}, Jeffrey I. Gordon⁵, Hal M. Hoffman⁹ & Richard A. Flavell^{1,8}

Non-alcoholic fatty liver disease (NAFLD) is the hepatic manifestation of metabolic syndrome and the leading cause of chronic liver disease in the Western world. Twenty per cent of NAFLD individuals develop chronic hepatic inflammation (non-alcoholic steatohepatitis, NASH) associated with cirrhosis, portal hypertension and hepatocellular carcinoma, yet the causes of progression from NAFLD to NASH remain obscure. Here, we show that the NLRP6 and NLRP3 inflammasomes and the effector protein IL-18 negatively regulate NAFLD/NASH progression, as well as multiple aspects of metabolic syndrome via modulation of the gut microbiota. Different mouse models reveal that inflammasome-deficiency-associated changes in the configuration of the gut microbiota are associated with exacerbated hepatic steatosis and inflammation through influx of TLR4 and TLR9 agonists into the portal circulation, leading to enhanced hepatic tumour-necrosis factor (TNF)- α expression that drives NASH progression. Furthermore, co-housing of inflammasome-deficient mice with wild-type mice results in exacerbation of hepatic steatosis and obesity. Thus, altered interactions between the gut microbiota and the host, produced by defective NLRP3 and NLRP6 inflammasome sensing, may govern the rate of progression of multiple metabolic syndrome-associated abnormalities, highlighting the central role of the microbiota in the pathogenesis of heretofore seemingly unrelated systemic auto-inflammatory and metabolic disorders.

The prevalence of non-alcoholic fatty liver disease (NAFLD) ranges from 20–30% in the general population and up to 75–100% in obese individuals^{1,2}. NAFLD is considered one of the manifestations of metabolic syndrome³. Whereas most patients with NAFLD remain asymptomatic, 20% progress to develop chronic hepatic inflammation (non-alcoholic steatohepatitis, NASH), which in turn can lead to cirrhosis, portal hypertension, hepatocellular carcinoma and increased mortality^{4–6}. Despite its high prevalence, factors leading to progression from NAFLD to NASH remain poorly understood and no treatment has proven effective^{7,8}.

A “two hit” mechanism is proposed to drive NAFLD/NASH pathogenesis⁹. The first hit, hepatic steatosis, is closely associated with lipotoxicity-induced mitochondrial abnormalities that sensitize the liver to additional pro-inflammatory insults. These second hits include enhanced lipid peroxidation and increased generation of reactive oxygen species (ROS)¹⁰. Inflammasomes are cytoplasmic multi-protein complexes composed of one of several NLR and PYHIN proteins, including NLRP1, NLRP3, NLRC4 and AIM2. Inflammasomes are sensors of endogenous or exogenous pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs)¹¹ that govern cleavage of effector pro-inflammatory cytokines such as pro-IL-1 β and pro-IL-18 (refs 12, 13). Most DAMPs trigger the generation of ROS, which are known to activate the NLRP3 inflammasome¹⁴. Therefore, we propose that inflammasome-dependent processing of IL-1 β and IL-18 may have an important role in the progression of NAFLD.

Results

Feeding adult mice a methionine-choline-deficient diet (MCDD) for 4 weeks beginning at 8 weeks of age induces several features of human NASH, including hepatic steatosis, inflammatory cell infiltration and ultimately fibrosis¹⁵. To investigate the role of inflammasomes in NASH progression, we fed MCDD to C57Bl/6 wild type (NCI), apoptosis-associated speck-like protein containing a CARD (*Asc*^{−/−}, also known as *Pycard*) and caspase 1 (*Casp1*^{−/−}) mutant mice to induce early liver damage in the absence of fibrosis (Fig. 1a–d and Supplementary Fig. 1c). Compared to wild-type animals, age- and gender-matched *Asc*^{−/−} and *Casp1*^{−/−} mice that were fed MCDD were characterized by significantly higher serum alanine aminotransferase (ALT) and aspartate aminotransferase (AST) activity, by enhanced microvesicular and macrovesicular hepatic steatosis, and by accumulation of multiple immune subsets in the liver from the innate and adaptive arms of the immune system (as defined by pathological examination and flow cytometry; $n = 7$ –11 mice per group; Fig. 1a–d and Supplementary Figs 1c, 2a). Remarkably, the hepatic accumulation of T and B cells seems to be dispensable for this phenotype because *Asc*^{−/−} mice lacking adaptive immune cells (*Asc*^{−/−}; *Rag*^{−/−}) also showed more severe NASH compared to wild-type animals, and comparable degrees of pathology to *Asc*^{−/−} animals (Supplementary Fig. 2b–d).

To test whether the increased NASH observed in *Asc*^{−/−} and *Casp1*^{−/−} deficient mice was mediated by IL-1 β or IL-18, we performed similar experiments using mice deficient in either the IL-1 receptor (*Il1r*^{−/−})

¹Department of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ²Department of Cell Biology, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ³Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ⁴Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ⁵Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St Louis, Missouri 63108, USA. ⁶Division of Allergy and Immunology, Department of Internal Medicine, Washington University School of Medicine, St Louis, Missouri 63108, USA. ⁷Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ⁸Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. ⁹Department of Pediatrics, Rady Children's Hospital San Diego, University of California at San Diego, La Jolla, California 92093, USA.

*These authors contributed equally to this work.

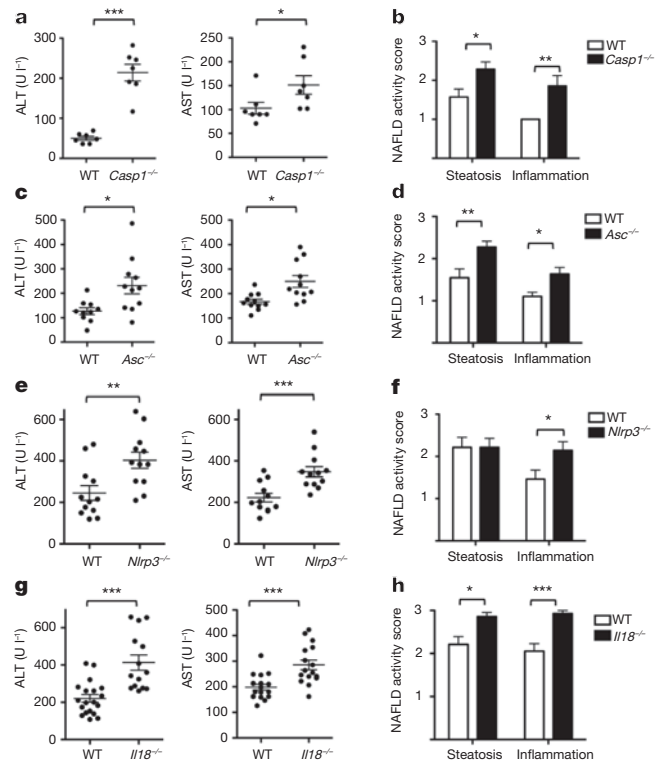


Figure 1 | Increased severity of NASH in inflammasome-deficient mice. To induce NASH, mice were fed with MCDD for 24 days. Their serum ALT and AST activities were measured and NAFLD histological activity scores were determined. **a–h**, Comparison of ALT, AST and NAFLD activity, plus histological scores for steatosis and inflammation between singly housed wild-type (WT) mice and *Casp1*^{−/−} (**a, b**), *Asc*^{−/−} (**c, d**), *Nlrp3*^{−/−} (**e, f**), or *Il18*^{−/−} (**g, h**). Data represent two independent experiments ($n = 7$ – 19 mice per treatment group). Error bars represent the s.e.m. of samples within a group. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ (Student's *t*-test).

or IL-18 (*Il18*^{−/−}). *Il1r*^{−/−} mice did not show any changes in the severity of NASH when compared to wild-type mice when fed MCDD (Supplementary Fig. 1a, b). In contrast to, but similar to *Asc*^{−/−} and *Casp1*^{−/−} mice, MCDD-fed *Il18*^{−/−} animals featured a significant exacerbation of NASH severity (Fig. 1g, h and Supplementary Fig. 1c).

To assess the role of the NLRP3 inflammasome in NASH progression, we fed singly housed *Nlrp3*^{−/−} and wild-type animals MCDD for 24 days and evaluated disease progression. *Nlrp3*^{−/−} mice developed exacerbated NASH compared to wild-type mice as judged by increased levels of serum ALT and AST, plus NAFLD activity inflammation scores (Fig. 1e, f and Supplementary Fig. 1c). Remarkably, bone marrow chimaeric mice in which NLRP3 and ASC deficiency was limited to the haematopoietic compartment did not show any increase in the severity of NASH when compared to wild-type mice reconstituted with wild-type bone marrow (Supplementary Fig. 3a–f). Likewise, knock-in mice that specifically express a constitutively active NLRP3 inflammasome in CD11c⁺ myeloid cells (*Nlrp3KI*; CD11c⁺-Cre) or hepatocytes (*Nlrp3KI*; albumin-Cre)¹⁶ did not feature any significant differences in MCDD-induced NASH severity as compared to wild-type mice (Supplementary Fig. 3g–l). These results indicate that aberrations in inflammasome function in cells other than hepatocytes or myeloid cells are key determinants of the enhanced disease progression in inflammasome-deficient mice.

We recently discovered that inflammasomes act as steady-state sensors and regulators of the colonic microbiota, and that a deficiency in components of two inflammasomes, NLRP6 (ref. 17) and NLRP3 (unpublished), both of which include ASC and caspase 1, and involve

IL-18 but not IL-1R, results in the development of an altered transmissible, colitogenic intestinal microbial community¹⁷. This microbiota is associated with increased representation of members of Bacteroidetes (Prevotellaceae) and the bacterial phylum TM7, and reductions in representation of members of the genus *Lactobacillus* in the Firmicutes phylum¹⁷. Moreover, electron microscopy studies disclosed aberrant colonization of crypts of Lieberkühn with bacteria with morphologic features of Prevotellaceae¹⁷. Therefore, we sought to investigate whether enhanced NASH severity in inflammasome-deficient mice is driven by their altered microbiota. Strikingly, co-housing of *Asc*^{−/−} and *Il18*^{−/−} mice with wild-type animals for 4 weeks (beginning at 4–6 weeks of age), before induction of NASH with MCDD resulted in significant exacerbation of NASH in the wild-type cage-mates (which we will refer to as WT(*Asc*^{−/−}) and WT(*Il18*^{−/−}), respectively, in the following text), as compared to singly housed, age- and gender-matched wild-type controls ($n = 5$ – 7 mice per genotype per housing condition). In co-housed wild-type mice, disease severity reached comparable levels to that of co-housed *Asc*^{−/−} and *Il18*^{−/−} mice (Fig. 2a–h). Moreover, significantly increased numbers of multiple inflammatory cell types were present in the liver of WT(*Asc*^{−/−}) compared to wild-type mice (Supplementary Fig. 2a). Similar findings were observed in wild-type mice co-housed with *Casp1*^{−/−}, *Nlrp3*^{−/−} and *Nlrp6*^{−/−} mice (Supplementary Fig. 4a–f). To exclude the possibility that aberrant microbiota presented in all mice maintained in our vivarium, we co-housed wild-type mice with other strains of NLR-deficient mice that were either obtained from the same source as *Asc*^{−/−} and *Nlrp3*^{−/−} mice (*Nlrp4*^{−/−}, *Nlrp12*^{−/−}), or generated in our laboratory (*Nlrp4c*^{−/−}). None of these strains featured a similar phenotype (Supplementary Fig. 4g–l). These results indicate that the transmissible colitogenic microbiota present in inflammasome-deficient mice is a major contributor to their enhanced NASH. In agreement with this, combined antibiotic treatment with ciprofloxacin and metronidazole, previously shown to abrogate the colitogenic activity

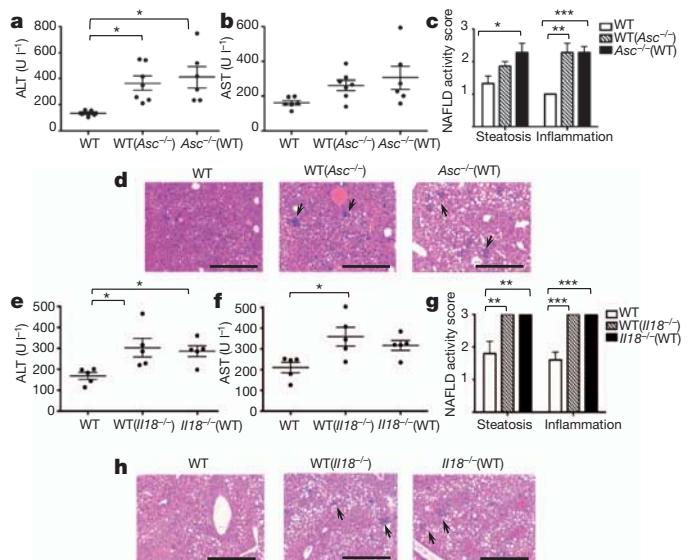


Figure 2 | Increased severity of NASH in *Asc*- and *Il18*-deficient mice is transmissible to co-housed wild-type animals. *Asc*^{−/−} or *Il18*^{−/−} mice and wild-type mice were co-housed for 4 weeks and then fed MCDD. **a–d**, ALT (**a**), AST (**b**), NAFLD activity scores (**c**), and haematoxylin and eosin-stained sections of livers (**d**) of singly housed wild-type mice (WT), wild-type mice co-housed with *Asc*^{−/−} mice (WT(*Asc*^{−/−})), and *Asc*^{−/−} mice co-housed with wild-type mice (*Asc*^{−/−}(WT)). **e–h**, ALT (**e**), AST (**f**), NAFLD activity histological scores (**g**), and haematoxylin and eosin-stained sections of livers (**h**) of wild-type, WT(*Il18*^{−/−}) and *Il18*^{−/−}(WT). Data are representative of two independent experiments. Error bars represent s.e.m. Scale bars, 200 μ m (**d, h**). * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

of the microbiota associated with inflammasome-deficient mice associated microbiota¹⁷, significantly reduced the severity of NASH in $Asc^{-/-}$ mice, and abolished transmission of the phenotype to $WT(Asc^{-/-})$ animals (Supplementary Fig. 5).

To ascertain the effects of MCDD on the gut microbiota, we performed a culture-independent analysis of amplicons generated by primers directed against variable region 2 of bacterial 16S ribosomal RNA genes of faecal samples collected from wild-type mice co-housed with $Asc^{-/-}$ animals ($WT(Asc^{-/-})$), their $Asc^{-/-}$ cage-mates ($Asc^{-/-}(WT)$) as well as singly housed wild-type controls 1 day and 12 days before, and 7, 14 and 19 days after initiation of this diet ($n = 20$ animals; 8 singly housed wild-type, 6 co-housed wild-type and 6 $Asc^{-/-}$ mice). The structures of bacterial communities were compared based on their phylogenetic content using unweighted UniFrac. The results are illustrated in Fig. 3. Supplementary Table 1 provides a list of all phylotypes that, based on criteria outlined in Methods, discriminate co-housed $WT(Asc^{-/-})$ from their singly housed wild-type counterparts. Prior to MCDD, and consistent with our previous findings¹⁷, the faecal microbiota of $WT(Asc^{-/-})$ mice adopted a configuration similar to $Asc^{-/-}$ cage-mates, including the appearance of Prevotellaceae (Supplementary Table 1 and Fig. 3 a–c). There was also a significant increase in proportional representation of members of the family Porphyromonadaceae (primarily in the genus *Parabacteroides*) in $WT(Asc^{-/-})$ mice compared to their singly

housed wild-type counterparts (Fig. 3d,e). The representation of Porphyromonadaceae was greatly increased in both the co-housed wild-type and $Asc^{-/-}$ mice (but not in singly housed wild-type) when they were switched to a MCDD diet ($P < 0.01$; t -test; Fig. 3d). A dramatic increase in the family Erysipelotrichaceae (phylum Firmicutes) also occurred with MCDD in both singly and co-housed WT animals, to a level that was $>10\%$ of the community (Fig. 3f). Although the Prevotellaceae decreased when co-housed $WT(Asc^{-/-})$ mice were placed on MCDD, their relative abundance remained significantly higher than in singly housed wild-type animals (Fig. 3c).

Together, these results pointed to the possibility that members of the altered intestinal microbiota in inflammasome-deficient MCDD-treated mice may promote a signalling cascade in the liver upon translocation, resulting in progression to NASH in susceptible animals. Toll-like receptors (TLR) have a major role in NAFLD pathophysiology due to the liver's exposure to relatively large amounts of PAMPs derived from the intestine and delivered via the portal circulation^{18–20}. Therefore, we propose that TLR signalling mediates the increased susceptibility to progression to NASH in mice exposed to the gut microbiota of $Asc^{-/-}$ animals. $Myd88^{-/-};Trif^{-/-}$ mice are devoid of all TLR signalling pathways. When co-housed with $Asc^{-/-}$ ($Myd88^{-/-};Trif^{-/-}(Asc^{-/-})$) mice between 5 and 9 weeks of age, they showed decreased severity of NASH after exposure to MCDD for 24 days, compared to $WT(Asc^{-/-})$ mice (Supplementary Fig. 6a, b).

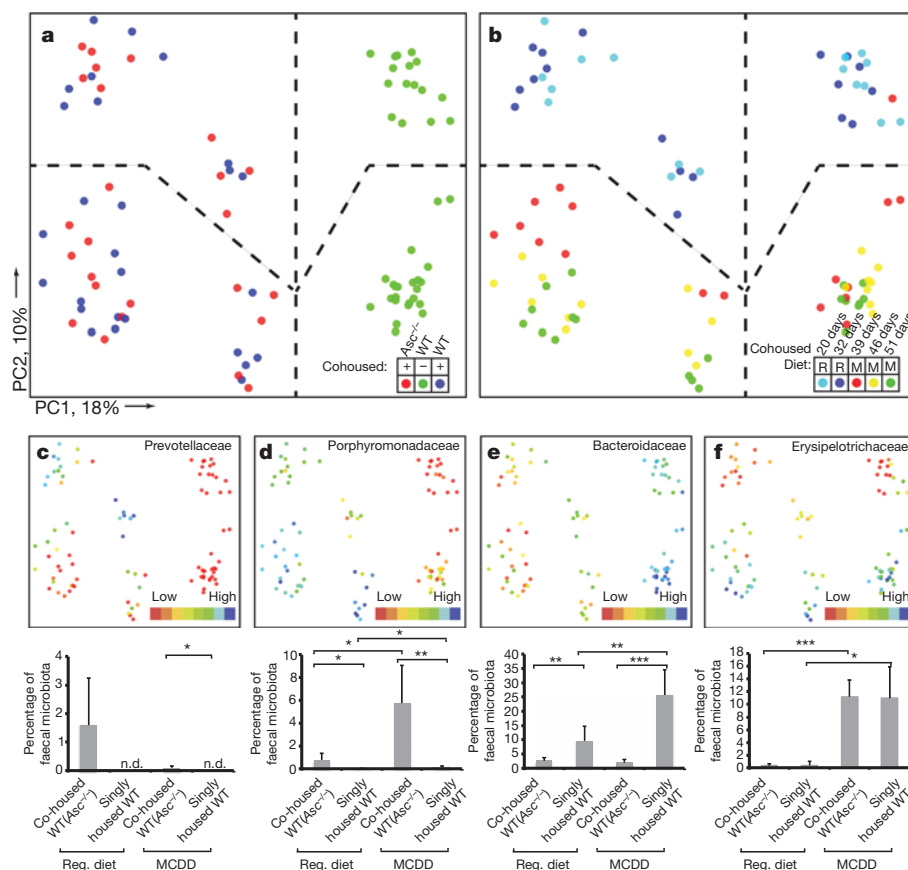


Figure 3 | 16S rRNA sequencing demonstrates diet and co-housing associated changes in gut microbial ecology. **a**, Principal coordinates analysis (PCoA) of unweighted UniFrac distances of 16S rRNA sequences demonstrating clustering according to co-housing status on principal coordinate 1 (PC1). **b**, PCoA of same plot as in **a** coloured for experimental day. Mice were co-housed and fed a regular diet (R) for the first 32 days of the experiment (two time points taken at day 20 and 32) before being switched to MCDD (M, sampled at days 39, 46 and 51 of the experiment). **c–f**, PCoA and bar graphs of family level taxa Prevotellaceae, Porphyromonadaceae, Bacteroidaceae and Erysipelotrichaceae demonstrating diet- and microbiota-

dependent differences in taxonomic representation. PCoA plots contain spheres representing a single faecal community coloured according to relative representation of the taxon (blue represents relatively higher levels; red indicates lower levels). Bar graphs represent averaged taxonomic representation for singly or co-housed mouse while on either regular or MCD diet ($n = 8$ for singly housed wild-type, $n = 12$ co-housed $Asc^{-/-}(WT)$ and $WT(Asc^{-/-})$ animals). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by t -test after Bonferroni correction for multiple hypotheses. n.d., not detected; Reg. diet, regular diet.

To define which specific TLRs were responsible for the inflammatory response, we co-housed *Tlr4*-, *Tlr9*- or *Tlr5*-deficient mice with *Asc*^{-/-} animals and induced NASH with MCDD as previously described. Similar to wild-type mice, *Tlr5*^{-/-} mice co-housed with *Asc*^{-/-} mice (*Tlr5*^{-/-}(*Asc*^{-/-})) featured a statistically significant exacerbation of hepatic injury, steatosis and inflammation, when compared to singly housed *Tlr5*^{-/-} controls (Fig. 4c and Supplementary Fig. 6g, h), indicating that TLR5 does not mediate the microbiota-mediated exacerbation in disease severity. In contrast, *Tlr4*^{-/-}(*Asc*^{-/-}) and *Tlr9*^{-/-}(*Asc*^{-/-}) mice did not show the customary increase in disease severity when compared to their singly housed *Tlr4*^{-/-} and *Tlr9*^{-/-} counterparts (Fig. 4a, b and Supplementary Fig. 6c–f).

These observations indicate that intact bacteria or bacterial products derived from the intestine trigger TLR4 and TLR9 activation, which results in an increased rate of disease progression in mice that house a colitogenic gut microbiota associated with inflammasome deficiency (that is, *Asc*^{-/-} and WT(*Asc*^{-/-}) mice). Efforts to sequence 16S rRNA genes that might be present in total liver DNA, microbial quantitative PCR assays of portal vein blood DNA, histologic analysis of intact

liver, and aerobic and anaerobic cultures of liver homogenates did not reveal any evidence of intact bacteria in wild-type or *Asc*^{-/-} mice fed MCDD (data not shown). Notably, transmission electron microscopy studies of colon collected from wild-type and *Asc*^{-/-} mice revealed an abundance of electron-dense material, suggestive of some black-pigmented bacterial species, in colonic epithelial cells and macrophages located in the lamina propria of *Asc*^{-/-} mice but not in wild-type animals (Fig. 4e and Supplementary Fig. 7c). In agreement with previous results, we did not detect any translocation of intact bacteria (Fig. 4e and Supplementary Fig. 7c).

These observations provide evidence for the uptake of bacterial products from locally invasive gut microbes in *Asc*^{-/-} mice (Fig. 4e and Supplementary Fig. 7c). If microbial components, rather than whole organisms, were transmitted to the liver then they should be detectable in the portal circulation. Indeed, levels of TLR4 and TLR9 agonists, but not TLR2 agonists (assayed by their ability to activate TLR reporter cell lines), were markedly increased in the portal circulation of MCDD-fed WT(*Asc*^{-/-}), and *Asc*^{-/-} mice compared to wild-type controls ($n = 13$ –28 mice per group; Fig. 4d and Supplementary Fig. 7a, b). Altogether, these results indicate a mechanism whereby TLR4 and TLR9 agonist efflux from the intestines of inflammasome-deficient mice or their co-housed partners, through the portal circulation, to the liver where they trigger TLR4 and TLR9 activation that in turn results in enhanced progression of NASH.

We next explored the downstream mechanism whereby microbiota-induced TLR signalling enhances NASH progression. Pro-inflammatory cytokines, and in particular TNF- α , a downstream cytokine of TLR signalling, are known to contribute to progression of hepatic steatosis to steatohepatitis and eventually hepatic fibrosis in a number of animal models and in human patients^{21,22}. Following induction of NASH by MCDD, hepatic *Tnf* mRNA expression was significantly upregulated in *Asc*^{-/-} and *Il18*^{-/-} mice, which show exacerbated disease, but not in *Il1r*^{-/-} mice, which do not (Supplementary Fig. 8a–c). Moreover, *Tnf* mRNA levels were significantly increased in wild-type mice that had been previously co-housed with *Asc*^{-/-} or *Il18*^{-/-} mice and then fed MCDD (Supplementary Fig. 8d, e), indicating that its enhanced expression was mediated by elements of the microbiota responsible for NASH exacerbation. In contrast, we did not observe any changes in *Il6* or *Il1b* mRNA levels in the livers of *Asc*^{-/-}, *Il18*^{-/-} or *Il1r*^{-/-} mice compared to wild-type controls (Supplementary Fig. 8a–c). Furthermore, whereas MCDD-administered singly housed *Tnf*^{-/-} mice had comparable NASH severity to singly housed wild-type animals (Fig. 4f–h and Supplementary Fig. 8f), co-housing with *Asc*-deficient mice before MCDD induction of NASH resulted in increased liver injury, hepatic steatosis and inflammation in wild-type mice but not in *Tnf*^{-/-} mice (Fig. 4f–h and Supplementary Fig. 8f). These results indicate that TNF- α mediates the hepatotoxic effects downstream of the transmissible gut microbiota present in *Asc*^{-/-} mice.

The aberrant gut microbiota in NLRP3 and NLRP6 inflammasome-deficient mice induces colonic inflammation through epithelial induction of CCL5 secretion¹⁷. To test whether this colon inflammation influences TLR agonist influx into the portal circulation and NASH progression, we induced NASH in wild-type and *Ccl5*^{-/-} mice that had been either singly housed or co-housed. MCDD-fed, singly housed wild-type and *Ccl5*^{-/-} mice showed equivalent levels of NASH severity (Supplementary Fig. 9a–c), indicating that CCL5 does not have a role in the early stages of NAFLD/NASH in the absence of the inflammasome-associated colitogenic microbiota. However, we documented significantly increased levels of liver injury, inflammation and steatosis in WT(*Asc*^{-/-}) but not *Ccl5*^{-/-}(*Asc*^{-/-}) mice (Fig. 5a–c), which led us to conclude that CCL5 is required for the exacerbation of disease through cohousing with inflammasome-deficient mice. Moreover, *Ccl5*^{-/-}(*Asc*^{-/-}) animals showed significantly reduced levels of TLR4 and TLR9 agonists in their portal vein blood than WT(*Asc*^{-/-}) mice (Supplementary Fig. 9d–f). Together, these results

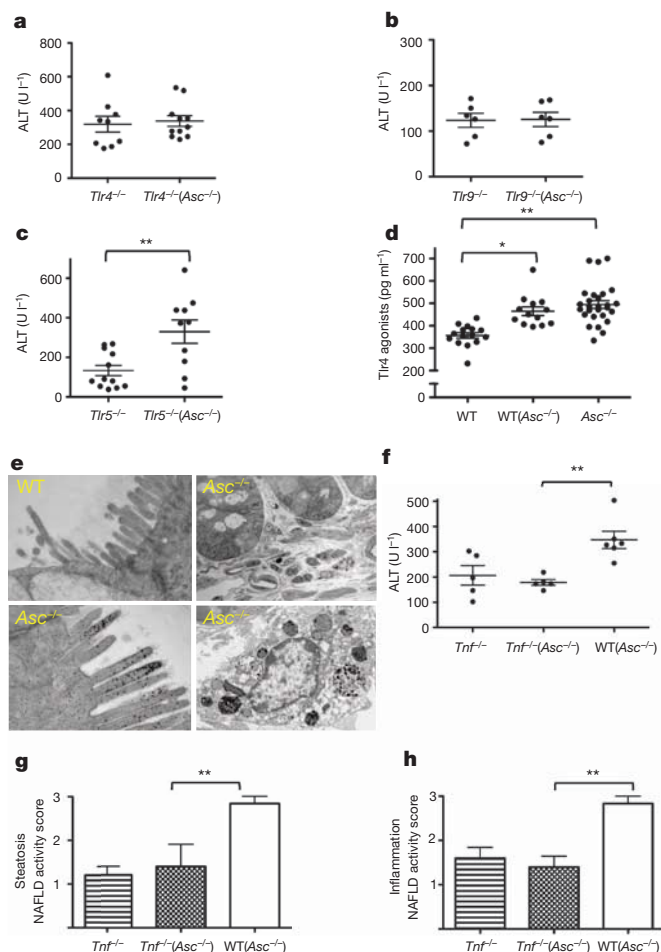


Figure 4 | Increased severity of NASH in *Asc*-deficient and co-housed wild-type animals is mediated by TLR4, TLR9 and TNF- α . *Asc*^{-/-} mice were co-housed with wild-type, *Tnf*^{-/-}, *Tlr4*^{-/-}, *Tlr9*^{-/-} or *Tlr5*^{-/-} mice for 4 weeks and then fed MCDD. **a–c**, ALT levels of *Tlr4*^{-/-}(*Asc*^{-/-}) (**a**), *Tlr9*^{-/-}(*Asc*^{-/-}) (**b**), and *Tlr5*^{-/-}(*Asc*^{-/-}) mice (**c**) and their singly housed counterparts. **d**, TLR4 agonists in portal vein sera from MCDD-fed wild-type, WT(*Asc*^{-/-}) and *Asc*^{-/-} animals. **e**, Transmission electron microscopy images of colon from wild-type and *Asc*^{-/-}. **f–h**, ALT (**f**) and NAFLD (**g–h**) activity histological scores of *Tnf*^{-/-}, WT(*Asc*^{-/-}) and *Tnf*^{-/-}(*Asc*^{-/-}) mice. Data are representative of two independent experiments. Error bars represent s.e.m. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

indicate that microbiota-induced subclinical colon inflammation is a determining factor in the rate of TLR agonist influx from the gut, and in NAFLD/NASH progression.

The MCDD system is a common model for studying inflammatory processes associated with progression from NAFLD to NASH, yet it lacks many of the associated metabolic phenotypes of NAFLD, such as obesity and insulin resistance²³. As such, our results in this model might conceivably be limited to the way dysbiosis can influence NASH progression in patients with enhanced intestinal permeability, such as those with inflammatory bowel disease²⁴, but not for the majority of patients who suffer from NASH in the context of metabolic syndrome. To test whether alterations in the gut microbiota of inflammasome-deficient mice may affect the rate of progression of NAFLD and other features associated with metabolic syndrome, we extended our studies to genetically obese mice and mice fed with high-fat diet (HFD).

Leptin-receptor deficient (*db/db*; *db* is also known as *Lepr*) animals develop multiple metabolic abnormalities, including NAFLD and impaired intestinal barrier function²⁵, that closely resemble the human disease²⁶. However, significant hepatocyte injury, inflammation, and fibrosis are not observed in the absence of a “second hit”²⁷. Upon co-housing of *db/db* mice with *Asc*^{-/-} (*db/db*(*Asc*^{-/-})) or WT mice (*db/db*(WT)) for a period of 12 weeks, and as previously shown for *Asc*^{-/-} mice¹⁷, the colon and ileum of all *db/db*(*Asc*^{-/-}) mice showed mild to moderate mucosal and crypt hyperplasia (Fig. 5d–f) that was not seen in *db/db*(WT) mice.

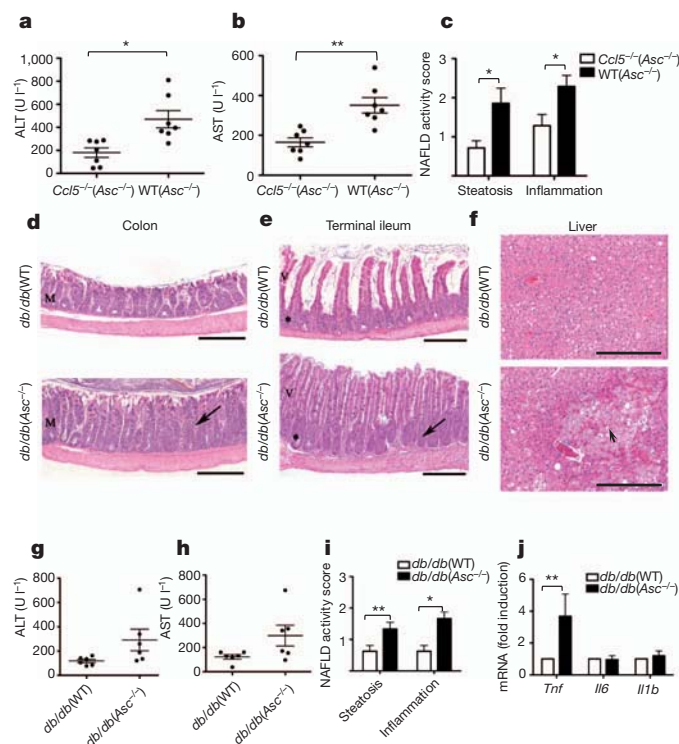


Figure 5 | Increased severity of NASH in *Asc*-deficient mice is transmissible to *db/db* by co-housing and is mediated by CCL5-induced intestinal inflammation. a–c, ALT (a), AST (b) and NAFLD (c) activity histological scores of WT(*Asc*^{-/-}) and *Ccl5*^{-/-}(*Asc*^{-/-}) mice. Data represents two independent experiments. d–j, *db/db* mice were co-housed with wild-type or *Asc*^{-/-} mice for 12 weeks. d–f, Representative haematoxylin and eosin-stained sections of colon (d), terminal ileum (e) and liver (f) from *db/db*(WT) and *db/db*(*Asc*^{-/-}) mice fed a standard chow diet. Mucosal and crypt hyperplasia (arrow). Hepatocyte degeneration (arrowhead). Scale bars, 500 μ m (d–e), 200 μ m (f). g–i, ALT (g), AST (h) and NAFLD (i) activity scores of *db/db*(WT) and *db/db*(*Asc*^{-/-}) mice. j, Hepatic *Tnf*, *Il6* and *Il1b* mRNA levels. Error bars represent s.e.m. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

Strikingly, co-housed *db/db*(*Asc*^{-/-}) mice also showed increased levels of hepatocyte injury as evidenced by higher levels of ALT and AST in their sera, and significantly exacerbated steatosis and hepatic inflammation scores when compared with *db/db*(WT) mice (Fig. 5g–i). In addition to a parenchymal inflammatory exudate, patchy areas of markedly degenerated hepatocytes and hepatocytes undergoing necrosis were observed, but only in *db/db*(*Asc*^{-/-}) animals (Fig. 5f). Furthermore, some areas of congestion were seen in the centro-lobular zone as well as in the hepatic parenchyma — features that resemble peliosis hepatis, a condition observed in a variety of pathological settings including infection (data not shown). In accord with our MCDD results, hepatic *Tnf* mRNA levels were significantly higher in co-housed *db/db*(*Asc*^{-/-}) mice than in *db/db*(WT) animals (Fig. 5j). Again, no significant differences were observed in hepatic *Il6* or *Il1b* mRNA levels (Fig. 5j).

Interestingly, *db/db*(*Asc*^{-/-}) mice developed significantly more weight gain compared to *db/db*(WT) mice after 12 weeks of co-housing (Fig. 6a), indicating that the inflammasome-associated gut microbiota could exacerbate additional processes associated with the metabolic syndrome, such as obesity. To address this possibility, we monitored multiple metabolic parameters in wild-type, WT(*Asc*^{-/-}) and *Asc*^{-/-} mice fed a high-fat diet (HFD) for 12 weeks. Strikingly, *Asc*^{-/-} mice gained body mass more rapidly and featured enhanced hepatic steatosis (Fig. 6b, c and Supplementary Fig. 11f). *Asc*^{-/-} mice also showed elevated fasting plasma glucose and insulin levels, and decreased glucose tolerance compared to singly housed weight-matched wild-type mice (Fig. 6d–f). Interestingly, WT(*Asc*^{-/-}) mice recapitulated

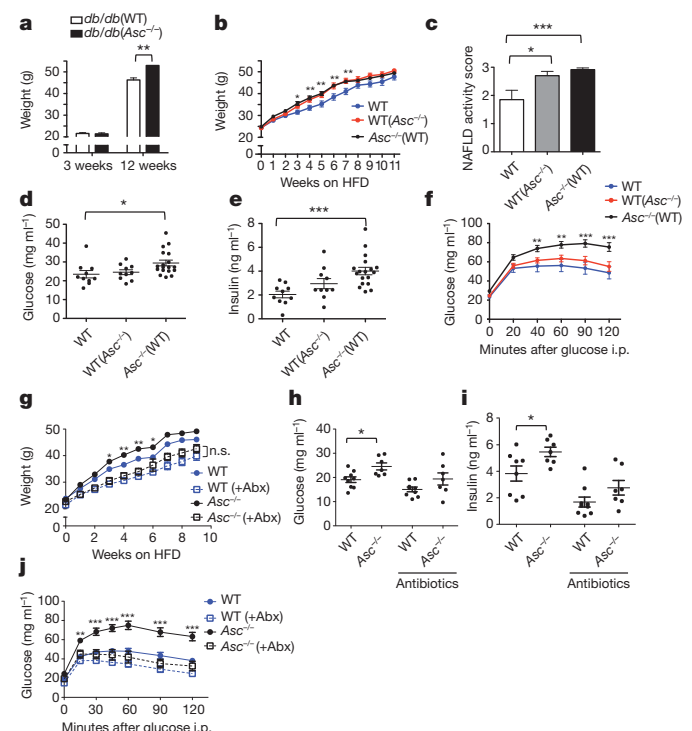


Figure 6 | *Asc*-deficient mice develop increased obesity and loss of glycaemic control on HFD. a, Weight of *db/db*(WT) or *db/db*(*Asc*^{-/-}) mice at 3 weeks of age and at 12 weeks of co-housing. b–f, *Asc*^{-/-} and wild-type mice were co-housed for 4 weeks and then fed HFD. b, Body weights. c, NAFLD histological activity score. d, e, Fasting plasma glucose and insulin after 11 weeks of HFD. f, Intraperitoneal (i.p.) glucose tolerance test after 12 weeks of HFD. g–j, Mice were untreated, or treated orally with antibiotics (Abx), for 3 weeks before HFD feeding for 12 weeks. g, Body weights. h, i, Fasting plasma glucose and insulin levels after 8 weeks on a HFD. j, Intraperitoneal glucose tolerance test after 10 weeks of HFD. Error bars represent s.e.m. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

the same increased rate of body mass gain and steatosis when compared to singly housed wild-type controls, although they did not show significant alterations in glucose homeostasis (Fig. 6d–f). Nevertheless, antibiotic treatment (ciprofloxacin and metronidazole) abrogated all these abnormalities, including altered rate of gain in body mass, glucose intolerance and fasting plasma insulin levels in *Asc*^{−/−} mice compared to wild-type mice (Fig. 6g–j). Alterations of these metabolic parameters were not caused by changes in feeding behaviour between the antibiotic-treated and untreated groups (data not shown). These results indicate different levels of microbiota-mediated regulation of the various manifestations of the metabolic syndrome: that is, some features (obesity, steatosis) are pronounced and transmissible by co-housing, whereas others (glycaemic control) are affected by alterations in the microbiota but not readily transferable by co-housing. Additionally, we performed a 16S rRNA-based analysis of the faecal microbiota of *Asc*^{−/−} and wild-type animals that were treated with or without ciprofloxacin and metronidazole (4 weeks) before switching to HFD for 4 additional weeks. Importantly, the analysis demonstrated that Prevotellaceae and Porphyromonadaceae, two family-level taxa, were undetectable in *Asc*^{−/−} mice 8 weeks after antibiotic treatment (Supplementary Fig. 12a–c; Supplementary Table 2).

To assess whether these metabolic abnormalities are specific to *Asc*^{−/−} mice, we performed similar experiments with *Nlr4*^{−/−} mice. These mice showed an equal rate of body mass gain, and similar glucose tolerance phenotypes as singly housed wild-type mice, confirming the specificity of the phenotype (Supplementary Fig. 10a–d). 16S rRNA analysis revealed that there was an increased representation of Porphyromonadaceae in *Nlr4*^{−/−} mice when compared to wild-type mice (Supplementary Table 3). These results indicate that (1) some metabolic aberrations associated with the dysbiosis of inflammasome-deficient mice can be horizontally transferred from one mouse to another, (2) the gut microbiota of inflammasome-deficient mice has a negative effect on NAFLD progression and glucose homeostasis, and (3) configurational changes in the microbiota, which involve overrepresentation Porphyromonadaceae in combination with alterations in additional taxa, are likely required to produce these host phenotypes.

Discussion

The results presented here provide evidence that modulation of the intestinal microbiota through multiple inflammasome components is a critical determinant of NAFLD/NASH progression as well as multiple other aspects of metabolic syndrome such as weight gain and glucose homeostasis. Our results demonstrate a complex and cooperative effect of two sensing protein families, namely NLRs and TLRs, in shaping metabolic events. In the gut, the combination of host-related factors such as genetic inflammasome deficiency-associated dysbiosis result in abnormal accumulation of bacterial products in the portal circulation. The liver, being a ‘first pass’ organ and thus exposed to the highest concentration of portal system products such as PAMPs, is expected to be most vulnerable to their effects, particularly when pre-conditioned by sub-clinical pathology such as lipid accumulation in hepatocytes. Indeed in our models, accumulation of TLR agonists was sufficient to drive progression of NAFLD/NASH even in genetically intact animals.

This ‘gut–liver axis’, driven by alterations in gut microbial ecology, may offer an explanation for a number of long-standing, albeit poorly understood, clinical associations. One example is the occurrence of primary sclerosing cholangitis (PSC) in patients with inflammatory bowel disease, particularly those with inflammation along the length of the colon. Coeliac disease, another inflammatory disorder with increased intestinal permeability, is associated with a variety of liver disorders, ranging from asymptomatic transaminasaemia, NAFLD, to primary biliary cirrhosis (PBC). In fully developed cirrhosis, complications associated with high mortality such as portal hypertension, variceal bleeding, spontaneous bacterial peritonitis and encephalopathy

are triggered by translocation of bacteria or bacterial components, providing another important example of the importance of the interplay between the microbiome, the immune response and liver pathology²⁸.

Recent reports suggest a complex role of inflammasome function in multiple manifestations of the metabolic syndrome. Activation of IL-1 β , mainly through cleavage by the NLRP3 inflammasome, promotes insulin resistance^{29,30}, atherosclerotic plaque formation³¹, and β cell death^{32,33}. Moreover, caspase-1 activation seems to direct adipocytes towards a more insulin-resistant phenotype³⁴. Conversely, *Il18*-deficient mice are prone to develop obesity, hyperphagia and insulin resistance³⁵. These discrepancies most probably reflect a hierarchical contribution of multiple inflammasome components in different metabolic processes, tissues and mouse models. In agreement with previous studies, we found increased obesity and insulin resistance in *Il18*-deficient mice fed with a HFD (data not shown). However, and in contrast to two previous reports^{30,34}, we showed that *Asc*^{−/−} mice are prone to obesity induction and hepatosteatosis, as well as impaired glucose homeostasis when fed a HFD. We propose that alterations in intestinal microbiota communities associated with multiple inflammasome deficiencies could account for these discrepancies and it should be added to the list of major environmental/host factors affecting manifestations and progression of metabolic syndrome in susceptible populations.

In the inflammasome-deficient setting, a significant expansion of Porphyromonadaceae was found following administration of MCDD and HFD, which was abolished by antibiotic treatment. Interestingly, one member of the family, *Porphyromonas*, has been associated with several components of the metabolic syndrome in both mice and humans, including atherosclerosis and diabetes mellitus^{36,37}. Moreover, expansion of this taxa is strongly associated with complications of chronic liver disease³⁶. More work is needed to further delineate the relevance of the suggested taxa discovered in our work to the pathogenesis and progression of human NAFLD/NASH and other features of the metabolic syndrome. Elucidation of similar or distinct mechanisms to the ones presented here, possibly linking Porphyromonadaceae expansion to a propensity for development of the metabolic syndrome, would be of importance to the field.

METHODS SUMMARY

Six- to eight-week-old male mice were fed a methionine-choline-deficient diet for 24 days. Eight- to ten-week-old male mice were fed a HFD ad libitum. This diet consists of 60% calories from fat and was administered for 10–12 weeks. Standard histology of liver, terminal ileum and colon were described previously¹⁷. The presence of immune cells in liver tissue was analysed by flow cytometry on livers digested with 0.5 mg ml^{−1} collagenase. Glucose tolerance test were performed after 10–12 weeks of consuming the HFD and mice were fasted overnight (~14 h), and injected intraperitoneally with D-glucose. Transmission electron microscopy was performed as previously described¹⁷. Data are expressed as mean \pm s.e.m. Differences were analysed by Student's *t*-test or ANOVA and post hoc analysis for multiple group comparison. *P* values \leq 0.05 were considered significant.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 April; accepted 22 December 2011.

Published online 1 February 2012.

- Sheth, S. G., Gordon, F. D. & Chopra, S. Nonalcoholic steatohepatitis. *Ann. Intern. Med.* **126**, 137–145 (1997).
- Ludwig, J., Viggiano, T. R., McGill, D. B. & Oh, B. J. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin. Proc.* **55**, 434–438 (1980).
- Marchesini, G. *et al.* Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology* **37**, 917–923 (2003).
- Caldwell, S. H. *et al.* Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. *Hepatology* **29**, 664–669 (1999).
- Shimada, M. *et al.* Hepatocellular carcinoma in patients with non-alcoholic steatohepatitis. *J. Hepatol.* **37**, 154–160 (2002).
- Propst, A., Propst, T., Judmaier, G. & Vogel, W. Prognosis in nonalcoholic steatohepatitis. *Gastroenterology* **108**, 1607 (1995).

7. Charlton, M. Cirrhosis and liver failure in nonalcoholic fatty liver disease: molehill or mountain? *Hepatology* **47**, 1431–1433 (2008).
8. Hjelkrem, M. C., Torres, D. M. & Harrison, S. A. Nonalcoholic fatty liver disease. *Minerva Med.* **99**, 583–593 (2008).
9. Day, C. P. & James, O. F. Steatohepatitis: a tale of two “hits”? *Gastroenterology* **114**, 842–845 (1998).
10. Sanyal, A. J. *et al.* Nonalcoholic steatohepatitis: association of insulin resistance and mitochondrial abnormalities. *Gastroenterology* **120**, 1183–1192 (2001).
11. Sutterwala, F. S., Ogura, Y. & Flavell, R. A. The inflammasome in pathogen recognition and inflammation. *J. Leukoc. Biol.* **82**, 259–264 (2007).
12. Martinon, F., Burns, K. & Tschopp, J. The inflammasome. *Mol. Cell* **10**, 417–426 (2002).
13. Agostini, L. *et al.* NALP3 forms an IL-1 β -processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. *Immunity* **20**, 319–325 (2004).
14. Zhou, R., Yazdi, A. S., Menu, P. & Tschopp, J. A role for mitochondria in NLRP3 inflammasome activation. *Nature* **469**, 221–225 (2011).
15. Varela-Rey, M. *et al.* Non-alcoholic steatohepatitis and animal models: understanding the human disease. *Int. J. Biochem. Cell Biol.* **41**, 969–976 (2009).
16. Brydges, S. D. *et al.* Inflammasome-mediated disease animal models reveal roles for innate but not adaptive immunity. *Immunity* **30**, 875–887 (2009).
17. Elinav, E. *et al.* NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* **145**, 745–757 (2011).
18. Rivera, C. A. *et al.* Toll-like receptor-4 signaling and Kupffer cells play pivotal roles in the pathogenesis of non-alcoholic steatohepatitis. *J. Hepatol.* **47**, 571–579 (2007).
19. Miura, K. *et al.* Toll-like receptor 9 promotes steatohepatitis by induction of interleukin-1 β in mice. *Gastroenterology* **139**, 323–334 e7 (2010).
20. Seki, E. *et al.* TLR4 enhances TGF- β signaling and hepatic fibrosis. *Nature Med.* **13**, 1324–1332 (2007).
21. Crespo, J. *et al.* Gene expression of tumor necrosis factor α and TNF-receptors, p55 and p75, in nonalcoholic steatohepatitis patients. *Hepatology* **34**, 1158–1163 (2001).
22. Li, Z. *et al.* Probiotics and antibodies to TNF inhibit inflammatory activity and improve nonalcoholic fatty liver disease. *Hepatology* **37**, 343–350 (2003).
23. Diehl, A. M. Lessons from animal models of NASH. *Hepatol. Res.* **33**, 138–144 (2005).
24. Broomé, U., Glaumann, H. & Hultcrantz, R. Liver histology and follow up of 68 patients with ulcerative colitis and normal liver function tests. *Gut* **31**, 468–472 (1990).
25. Guo, X. *et al.* Leptin signaling in intestinal epithelium mediates resistance to enteric infection by *Entamoeba histolytica*. *Mucosal Immunol.* **4**, 294–303 (2011).
26. Ikejima, K. *et al.* The role of leptin in progression of non-alcoholic fatty liver disease. *Hepatol. Res.* **33**, 151–154 (2005).
27. Guebre-Xabier, M. *et al.* Altered hepatic lymphocyte subpopulations in obesity-related murine fatty livers: potential mechanism for sensitization to liver damage. *Hepatology* **31**, 633–640 (2000).
28. Almeida, J., Galhenage, S., Yu, J., Kurtovic, J. & Riordan, S. M. Gut flora and bacterial translocation in chronic liver disease. *World J. Gastroenterol.* **12**, 1493–1502 (2006).
29. Vandanmagsar, B. *et al.* The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. *Nature Med.* **17**, 179–188 (2011).
30. Wen, H. *et al.* Fatty acid-induced NLRP3-ASC inflammasome activation interferes with insulin signaling. *Nature Immunol.* **12**, 408–415 (2011).
31. Duewell, P. *et al.* NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature* **464**, 1357–1361 (2010).
32. Zhou, R., Tardivel, A., Thorens, B., Choi, I. & Tschopp, J. Thioredoxin-interacting protein links oxidative stress to inflammasome activation. *Nature Immunol.* **11**, 136–140 (2010).
33. Masters, S. L. *et al.* Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1 β in type 2 diabetes. *Nature Immunol.* **11**, 897–904 (2010).
34. Stienstra, R. *et al.* Inflammasome is a central player in the induction of obesity and insulin resistance. *Proc. Natl Acad. Sci. USA* **108**, 15324–15329 (2011).
35. Netea, M. G. *et al.* Deficiency of interleukin-18 in mice leads to hyperphagia, obesity and insulin resistance. *Nature Med.* **12**, 650–656 (2006).
36. Bajaj, J. S. *et al.* Linkage of gut microbiome with cognition in hepatic encephalopathy. *Am. J. Physiol. Gastrointest. Liver Physiol.* **302**, 168–175 (2011).
37. Makiura, N. *et al.* Relationship of *Porphyromonas gingivalis* with glycemic level in patients with type 2 diabetes following periodontal treatment. *Oral Microbiol. Immunol.* **23**, 348–351 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Eynon, J. Alderman, A. Williams, F. Manzo and H. Elinav for technical assistance and discussions; M. Graham and C. Rahner for performing electron microscopy; D. R. Peaper for assistance in microbiological culture procedures; R. Sherwin for helpful advice; X. Fan for technical assistance; Yale Diabetes Endocrinology Research Center and Mouse Metabolic Phenotyping Center for assistance with the metabolic analysis. E.E. is supported by the Cancer Research Institute (2010–2012) and by a supplementary grant from the Israel-US educational foundation (2009) and is a recipient of the Claire and Emmanuel G. Rosenblatt award from the American Physicians for Medicine in Israel Foundation (2010–2011). J.H.M. and T.S. are supported by Leukemia and Lymphoma Society Postdoctoral Fellowships. S.C.E. is supported by T32HL007974 and K08A1085038. W.Z.M. is supported by R01DK076674-01 and the VA Merit award. This work was supported in part by the Howard Hughes Medical Institute (G.I.S., R.A.F.), the United States-Israel binational Foundation grant (E.E. and R.A.F.), the Crohn's and Colitis Foundation of America (A.K. and J.I.G.) and R01 DK-40936, R24 DK-085638, P30 DK-45735 and U24 DK-059635. The authors report no conflict of interest.

Author Contributions J.H.-M., E.E. and R.A.F. designed the study and wrote the manuscript. J.H.-M., E.E., C.J., L.H., W.Z.M., M.J.J., J.-P.C., G.I.S. and C.A.T. performed the *in vitro* and *in vivo* experimental work and edited the manuscript. T.S. and S.C.E. supported the work with key suggestions and editing of the manuscript. H.M.H. provided the *Nlrp3* knock-in mice and provided valuable feedback on the manuscript. A.L.K. and J.I.G. performed the stool processing and metagenomic analysis of the microbiota and provided key suggestions to the manuscript and participated in its editing. R.A.F. directed the project.

Author Information 16S rRNA data sets have been deposited in MG-RAST under accession number qiime:909. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.A.F. (richard.flavell@yale.edu).

METHODS

Mice. *Casp1*^{tm1Flv} (*Casp1*^{tm1Flv}) and *Nlrp4c*^{-/-} mice were generated in our laboratory³⁸. Production of *ASC*^{-/-} (*Pycard*^{tm1Flv}), *Nlrp3*^{-/-}, *Nlrp6*^{-/-}, *Nlrp4*^{-/-} and *Nlrp12*^{-/-} mice is described elsewhere¹⁷. *Il18*^{-/-} (*Il18*^{tm1Aki}), *Il1r*^{-/-} (*Il1r1*^{tm1Imx}), *Tnf*^{-/-} (*Tnf*^{tm1Gkl}), *Thr4*^{-/-} (*Thr4*^{lps-del}), *Thr5*^{-/-} (*Thr5*^{tm1Flv}), *Myd88*^{-/-} (*Myd88*^{tm1Defr}), *Ccl5*^{-/-} (*Ccl5*^{tm1Hso}), *Rag1*^{-/-} (*Rag1*^{tm1Mom}), *CD11c-Cre* (*Itgax-cre*), *albumin-Cre* (*Alb-cre*), *Trif*^{-/-} (*Ticam1*^{Lps2}) and *db/db* (*Lep*^{db}) mice were obtained from Jackson Laboratories. *Thr9*^{-/-} mice have been described in another report³⁹. Production of *Nlrp3KI* (A350V) mice is described elsewhere¹⁶. Wild-type C57Bl/6 mice were purchased from the NCI. For co-housing experiments, age-matched wild-type and KO mice at the age of 4–6 weeks were co-housed in sterilized cages for 4 or 12 weeks at a ratio of 1:1 (WT:KO), with unrestricted access to food and water. No more than 6 mice in total were housed per cage. For antibiotic treatment, mice were given a combination of ciprofloxacin (0.2 g l⁻¹) and metronidazole (1 g l⁻¹) for 4 weeks in the drinking water. All antibiotics were obtained from Sigma Aldrich. All experimental procedures were approved by the local IACUC.

NASH model. 6–8 week-old male mice were fed a methionine-choline-deficient diet (MP Biomedicals) for 24 days. Methionine-choline-sufficient control diet was the same but supplemented with choline chloride (2 g per kg of diet) and DL-methionine (3 g per kg of diet). Mice had unrestricted access to food and water.

High fat diet model. 8–10 week-old male mice were fed a HFD ad libitum. This diet consists of 60% calories from fat (D12492i; Research Diets) and was administered for 10–12 weeks.

Histology. The intact liver was excised immediately after mice were euthanized by asphyxiation, fixed in 10% neutral buffered formalin and embedded in paraffin. Liver sections were stained with haematoxylin and eosin, or trichrome. Histological examination was performed in a blinded fashion by an experienced gastrointestinal pathologist with the histological scoring system for NAFLD⁴⁰. Briefly, steatosis and inflammation scores ranged from 0 to 3 with 0 being within normal limits and 3 being most severe. Individual scores were assigned for each parameter. The most severe area of hepatic inflammation of representative histology sections were photographed using an Olympus microscope.

Colons were fixed in Bouin's medium and embedded in paraffin. Blocks were serially sectioned along the cephalocaudal axis of the gut to the level of the lumen; 5-µm-thick sections were stained with haematoxylin and eosin. Digital light microscopic images were recorded with a Zeiss Axio Imager.A1 microscope, AxioCam MRC5 camera and AxioVision 4.7.1 imaging software (Carl Zeiss Microimaging). Further details in ref. 17.

Gene expression analysis. Tissues were preserved in RNAlater solution (Ambion), and subsequently homogenized in TRIzol reagent (Invitrogen). RNA (1 µg) was used to generate complementary DNA using the HighCapacity cDNA Reverse Transcription kit (Applied Biosystems). Real time PCR was performed using gene-specific primer/probe sets (Applied Biosystems) and Kapa Probe Fast qPCR kit (Kapa Biosystems) on a 7500 Fast Real Time PCR instrument (Applied Biosystems). The reaction conditions were 95 °C for 20 s, followed by 40 cycles of 95 °C for 3 s and 60 °C for 30 s. Data was analysed using the Sequence Detection Software according to the ΔC_t method with *Hprt* serving as the reference housekeeping gene.

Glucose tolerance test (GTT). GTTs were performed after 10–12 weeks of consuming the HFD. Mice were fasted overnight (~14 h), and injected intraperitoneally with 10% dextrose at a dose of 1 g per kg body weight. Blood was collected from tail vein and plasma glucose levels measured at indicated times using a YSI 2700 Select Glucose Analyzer (YSI Life Sciences). Plasma insulin levels were determined by radioimmunoassay (Linco).

Flow cytometry analysis. Livers were collected, digested with 0.5 mg ml⁻¹ collagenase IV (Sigma) for 45 min at 37 °C, homogenized and repeatedly centrifuged at 400g for 5 min to enrich for haematopoietic cells. Cells were stained for flow cytometry using antibodies against CD45.2, CD11b, CD11c, NK1.1, B220, CD4, CD8, TCRβ, F4/80, Gr-1, MHC class II (Biolegend) and analysed on a BD LDR II.

Portal vein blood collection. Mice were anaesthetized with ketamine 100 mg per kg and xylazine 10 mg per kg. Mice were placed on a clean surgical field, and the abdominal fur was clipped and cleaned with a two stage surgical scrub consisting of Betadine and 70% ethanol. A 1 to 1.5 cm midline incision was made in the skin and abdominal wall. The peritoneum was moved to the left and the portal vein was punctured with a 30G needle. Between 0.2 and 0.3 ml of blood were collected per mouse. Serum was recovered by centrifugation at 1,500g for 15 min at room temperature and then stored at -80 °C in endotoxin-free tubes until assayed.

Measurement of PAMPs. TLR2, TLR4 and TLR9 agonists were assayed in portal vein serum using HEK-blue mTLR2, HEK-blue mTLR4 and HEK-blue mTLR9

reporter cell lines (InvivoGen) and the manufacturer's protocol with modifications. In brief, 2.2 × 10⁵ HEK-blue mTLR2, 1.0 × 10⁵ HEK-blue mTLR4 and 2.0 × 10⁵ HEK-blue mTLR9 cells were plated in 96-well plates containing 10 µl of heat-inactivated (45 min at 56 °C) portal vein serum. Cells were then incubated for 21 h at 37 °C under an atmosphere of 5% CO₂/95% air. Twenty microlitres of the cell culture supernatants were collected and added to 180 µl of the QUANTI-Blue substrate in a 96-well plate. The mixtures were then incubated at 37 °C in 5% CO₂/95% air for 3 h and secreted embryonic alkaline phosphatase levels were determined using a spectrophotometer at 655 nm.

Transmission electron microscopy. Mice were perfused via their left ventricles using 4% paraformaldehyde in PBS. Selected tissues were fixed in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.4 for 1–2 h. Samples were rinsed three times in sodium cacodylate buffer and post-fixed in 1% osmium tetroxide for 1 h, en bloc stained in 2% uranyl acetate in maleate buffer pH 5.2 for a further hour then rinsed, dehydrated, infiltrated with Epon812 resin, and baked overnight at 60 °C. Hardened blocks were cut using a Leica UltraCut UCT. 60-nm-thick sections were collected and stained using 2% uranyl acetate and lead citrate. Samples were all viewed in an FEI Tencai Biotwin TEM at 80 kV. Images were taken using Morada CCD and iTEM (Olympus) software. Further details in ref. 17.

Bone marrow chimeras. Bone marrow was flushed from femurs with DMEM with 10% FBS, red cells were lysed, and the material filtered through a 70 µm filter. 10⁶ cells in 100 µl PBS were delivered by retro-orbital injection into lethally irradiated (1,000 rad) mice. For 2 weeks post-engraftment, mice were maintained on antibiotics (Sulfatrim). Six weeks after transplantation animals were switched to MCDD. A wild-type non-irradiated mouse was co-housed with the engrafted mice for 4 weeks before NASH induction. Under our standardized protocol, bone marrow chimeras routinely show a level of engraftment of ≥ 93%.

Bacterial 16S rRNA amplicon sequencing. Total DNA was isolated from the livers of mice fed a MCDD diet and used for attempted PCR amplification of variable region 2 of bacterial 16S rRNA genes²⁰ that may be present in the tissue. Thirty cycles of amplification of liver DNA prepared from seven wild-type, and seven *Asc*^{-/-} mice yielded detectable product (>60 ng per reaction) in three samples from the wild-type group and three samples from the *Asc*^{-/-} group. All amplicons were then subjected to multiplex pyrosequencing with a 454 instrument using FLX Titanium chemistry (137–1,510 reads per sample, average read length, 360 nucleotides). Reads were analysed using the QIIME software package. Operational taxonomic unit (OTU) picking was performed using uclust and taxonomic assignments made with RDP⁴¹. This analysis demonstrated inconsistent representation of taxa between animals and taxa that largely represented organisms not associated with the gut microbiota. G-test indicated that there was no significant correlation between any of these taxa and the presence of NASH.

For analysis of the faecal microbiota of MCDD-fed *Asc*^{-/-} (WT), WT(*Asc*^{-/-}) and singly housed wild-type mice, faecal pellets were collected at the time points indicated in Fig. 3. The protocols that we used to extract faecal DNA and to perform multiplex pyrosequencing of amplicons generated by PCR from the V2 regions of bacterial 16S rRNA genes, have been previously described²⁰. A total of 366,283 sequences were generated from 181 faecal samples (average 2,023 ± 685 reads per sample; average read length, 360 nucleotides). Sequences were de-multiplexed and binned into species-level operational taxonomic units (OTUs; 97% nucleotide sequence identity; %ID) using QIIME 1.2.1 (ref. 41). Taxonomy was assigned within QIIME using RDP. Chimaeric sequences were removed using ChimeraSlayer and OTUs were filtered to a minimum of 10 sequences per OTU and 1,000 OTUs per sample. PCoA plots were generated by averaging the unweighted UniFrac distances of 100 subsampled OTU tables. Statistical analysis was performed on the proportional representation of taxa (summarized to Phyla, Class, Order, Family and Genus levels), using paired (where possible) and unpaired *t*-tests. Taxa that were significantly different after multiple hypothesis testing were included in Supplementary Tables 1–3.

Statistical analysis. Data are expressed as mean ± s.e.m. Differences were analysed by Student's *t*-test or ANOVA and post hoc analysis for multiple group comparison. *P* values ≤ 0.05 were considered significant.

38. Sutterwala, F. S. *et al.* Critical role for NALP3/CIA1/Cryopyrin in innate and adaptive immunity through its regulation of caspase-1. *Immunity* **24**, 317–327 (2006).
39. Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
40. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
41. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).